

ANDERSON FONSECA
LUIZ PAULO DA CAS CICHOSKI
ORGANIZADORES



AS COLÔNIAS
DE FORMIGAS
SÃO CONSCIENTES?

MENTES COLETIVAS E IMPLICAÇÕES FILOSÓFICAS

**AS COLÔNIAS DE FORMIGAS SÃO CONSCIENTES?
MENTES COLETIVAS E IMPLICAÇÕES FILOSÓFICAS**

*ARE ANT COLONIES CONSCIOUS?
COLLECTIVE MINDS AND PHILOSOPHICAL IMPLICATIONS*

Comitê Editorial

André Leclerc

Universidade de Brasília

Daniel De Luca Noronha

Faculdade Jesuíta de Filosofia e Teologia

Eros Moreira de Carvalho

Universidade Federal do Rio Grande do Sul e bolsista de Produtividade em Pesquisa do CNPq

Giovanni Rolla

Universidade Federal da Bahia

Gustavo Leal Toledo

Universidade Federal de São João del-Rei

Marco Aurélio Sousa Alves

Universidade Federal de São João del-Rei

Oswaldo Frota Pessoa Junior

Universidade de São Paulo

Raquel Albieri Krempel

Universidade Federal do ABC

Rodrigo Azevedo dos Santos Gouvêa

Universidade Federal do Rio de Janeiro

Sérgio Farias de Souza Filho

Universidade Federal Rural de Pernambuco

**AS COLÔNIAS DE FORMIGAS SÃO CONSCIENTES?
MENTES COLETIVAS E IMPLICAÇÕES FILOSÓFICAS**

*ARE ANT COLONIES CONSCIOUS?
COLLECTIVE MINDS AND PHILOSOPHICAL IMPLICATIONS*

Organizadores

Anderson Fonseca

Luiz Paulo Da Cas Cichoski



Diagramação: Marcelo Alves

Capa: Gabrielle do Carmo

Revisão: Ana Luisa Lima Grein



A Editora Fi segue orientação da política de distribuição e compartilhamento da Creative Commons Atribuição-Compartilhamento 4.0 Internacional https://creativecommons.org/licenses/by/4.0/deed.pt_BR



O padrão ortográfico e o sistema de citações e referências bibliográficas são prerrogativas de cada autor. Da mesma forma, o conteúdo de cada capítulo é de inteira e exclusiva responsabilidade de seu respectivo autor.

Dados Internacionais de Catalogação na Publicação (CIP)

C719 As colônias de formigas são conscientes?: Mentes coletivas e implicações filosóficas [recurso eletrônico] / Anderson Fonseca e Luiz Paulo Da Cas Cichoski (org.). – Cachoeirinha : Fi, 2025.

335p.

Contém textos em inglês.

ISBN 978-65-5272-128-0

DOI 10.22350/9786552721280

Disponível em: <http://www.editorafi.org>

1. Filosofia – Mentes coletivas. I. Fonseca, Anderson. II. Cichoski, Luiz Paulo Da Cas.

CDU 130.2

SUMÁRIO

Prefácio	7
<i>João de Fernandes Teixeira</i>	
Apresentação	9
<i>Anderson Fonseca</i>	
<i>Luiz Paulo Cichoski</i>	
1	19
É a colônia de formigas um organismo consciente?	
<i>Anderson Luiz do Vale Fonseca</i>	
2	53
What is it like to be an ant colony?	
<i>Luke Roelofs</i>	
3	87
Towards a science of consciousness and social complexity ... for ants	
<i>Maria Luiza Iennaco</i>	
<i>Daniel Friedman</i>	
4	115
A memória imunológica composta por formigas	
<i>Aryel C. Goes</i>	
5	147
Memória coletiva e consciência coletiva são fenômenos científicos?	
<i>José Carlos Camillo</i>	
6	175
Ontologia social para insetos eussociais	
<i>Valdenor Monteiro Brito Júnior</i>	
7	211
Como ser realista sobre mente de grupos?	
<i>Luiz Paulo Da Cas Cichoski</i>	

8

261

Mentes de grupos e tipos naturais

Robert Rupert

9

309

Introspecção em mentes de grupos, desunião da consciência e pessoas indiscretas

Eric Schwitzgebel

Sophie R. Nelson

PREFÁCIO

*João de Fernandes Teixeira*¹

Prefaciando um livro que tem como tema central a existência de mentes coletivas é uma tarefa desafiadora. Essa é, contudo, a proposta de seus organizadores que conseguiram atrair um punhado de pesquisadores altamente qualificados para discutir esse problema nos mais variados aspectos.

A tentação de comparar sociedades humanas a formigueiros é uma metáfora recorrente. Tudo se passa como se minúsculos cérebros estivessem reunidos num cérebro maior. A metáfora aparece num dos livros de Daniel Dennett (*Brainstorms*, 1978) na qual ele alude à ideia de que a inteligência surge a partir de uma multiplicidade de elementos que não são inteligentes. É relativamente fácil validar essa metáfora para o cérebro humano. Não há neurônios inteligentes, mas sua conexão tem como resultado padrões de comportamento inteligente. Mas, para pensar dessa maneira, temos de lançar mão da ideia controversa de propriedade emergente, que ainda encontra muita resistência na filosofia da mente.

O próprio Dennett parece ter mudado de ideia sobre esse tema. Numa discussão com Chalmers ele argumenta que não faz sentido usar o termo “saúde” como um conceito suplementar ao fato de que todos os órgãos de um ser vivo estão em funcionamento. Quem não reconheceria, nessas críticas, a herança de Gilbert Ryle, de quem Dennett foi discípulo. Numa passagem clássica da literatura sobre filosofia da mente, Ryle leva a passear um amigo na Universidade de Oxford e, depois de avistar

¹ Professor aposentado da Universidade Federal de São Carlos

vários prédios, esse amigo reclama, para sua surpresa que ele não tinha visto nenhuma universidade.

Ned Block, um filósofo da geração de Dennett afirmou que se para cada habitante da China fosse fornecida uma bandeira e eles a brandissem simultaneamente a simulação de um cérebro poderia surgir. O argumento de Block só tangencia o problema abordado por Ryle e Dennett, que visavam desconstruir a ideia de propriedade emergente. Block queria atacar o funcionalismo, mas acaba fornecendo um ataque oblíquo à ideia de propriedade emergente.

Esse é apenas o pano de fundo das discussões filosóficas envolvidas nos textos dos autores deste livro. Será que a ideia de mente coletiva exige a validação do conceito de propriedade emergente? E se esse conceito for validado e aplicado, o que aconteceria quando uma mente coletiva fosse decomposta? Como localizar, em cada formiga, as características que a permitem integrar a mente coletiva à qual ela pertence?

APRESENTAÇÃO

Anderson Fonseca

Luiz Paulo Cichoski

As formigas estão entre as formas de vida animal mais abundantes do planeta, alcançando a marca de 2.10^{16} a 10^{24} indivíduos, superando em número todo o conjunto de seres terrestres, incluindo o homem. Entretanto, seu aspecto mais fascinante está em seu comportamento social, o qual fundamenta a colônia. Podemos pensar as formigas e sua relação com as colônias que formam como a interação de minúsculos cérebros interagindo entre si formando um (macro)cérebro, como se fossem os neurônios de um cérebro espacialmente distribuído. A relação entre as formigas é coerente o suficiente para gerar um ‘corpo’ integrado com uma fronteira física definida – ainda que epistemicamente indeterminada. Essas características da colônia de formigas motivam um estudioso do problema da consciência a se perguntar se não seria ela uma entidade consciente. Esse ponto desafia nossa noção de consciência. Ora, se um sistema com corpos individuais espalhados, como uma colônia de formigas, pode ter algum tipo de mente unificada, então, certos coletivos formados por organismos com sistemas nervosos simples ou complexos podem produzir conjuntamente uma consciência.

Este contexto nos leva à questão fulcral deste livro: É possível que coletivos tenham uma mente? Mais especificamente, a hipótese de que uma colônia de formigas seja uma entidade consciente é razoável? Se a resposta for positiva, quais seriam as implicações filosóficas? Talvez, uma delas seria a afirmação de que todo organismo formado pela reunião de microrganismos e células com um nível mediano de integração é um coletivo consciente. Nesse sentido, o ser humano seria

um *coletivo* consciente. Outra implicação seria que grupos humanos, cidades, torcidas organizadas, ou, até mesmo certos países como os EUA, o Brasil e a China são conscientes. Pode-se a partir daí também afirmar que organizações de insetos eussociais, como abelhas e cupins, são conscientes, incluindo, quem sabe, bactérias e minúsculos animais marinhos coloniais. Até mesmo grupos formados por agentes de Inteligência Artificial poderiam compor uma mente coletiva consciente. Isso significa que coletivos de seres inteligentes são modelos gradativos de consciência e podem teoricamente contribuir para o entendimento de como se forma a mente do animal humano. Essa perspectiva mais biológica e naturalista se configura como um novo programa de pesquisa que expande a proposta das Ciências Cognitivas de que sistemas cognitivos são constituídos por estruturas distribuídas.

A questão sobre a possibilidade de mentes coletivas, é claro, pode ser respondida por diferentes perspectivas: a visão reducionista, de que a consciência do todo pode ser explicada pela consciência das partes; ou, o ponto de vista estruturalista, de que a ordem como estão as partes determina a mente do todo; ou, a emergentista, de que a interação dos componentes produz uma mente *acima* e distinta do conjunto de membros; no entanto, cada perspectiva se fundamenta em uma definição própria de *consciência*. Há ainda a ideia negativa, de que nenhuma mente adicional seja produzida e o que consideramos ser um sistema cognitivo complexo, nada mais é do que um punhado de relações vazias de informação sofisticada. Mas analisar se uma colônia de formigas é um todo consciente nos coloca diante da ideia de que haja mentes fracionadas (i.e., que são a fração de um todo mental), mentes espalhadas, e, talvez, mentes intermediárias. Essas ideias, evidentemente, podem ser *estranhas*, porém, contribuem para a investigação dos mistérios da consciência.

Posto isto, ao olharmos para o minúsculo universo das formigas nos maravilhamos com a possibilidade de uma mente emergir de sua emaranhada teia de conexões e que estruturas similares possam também ser sistemas cognitivos. Por essa razão este livro foi organizado com o objetivo de reunir diferentes pontos de vista sobre esse tema, apresentando uma visão ampla e robusta das investigações sobre Consciência Coletiva. Grande parte das contribuições seguirá a questão central do livro, apresentando uma abordagem naturalista, biologicamente centrada, sobre Consciência Coletiva. No entanto, alguns dos textos trarão perspectivas mais tradicionais, focados mais nas experiências humanas, tendo, portanto, uma abordagem mais antropológica do que zoológica. O leitor também encontrará uma diversidade de posicionamentos sobre nossa questão central. Ainda que a maioria dos autores esboce simpatia pela tese da Consciência Coletiva, buscamos incluir vozes mais hesitantes e até mesmo contrárias a ela.

O presente livro também pretende estabelecer a linha de estudo sobre a Consciência Coletiva no Brasil, área que investiga a possibilidade de grupos de indivíduos inteligentes constituírem mentes adicionais. A partir desta obra, esperamos estimular outros pesquisadores a investigarem essa linha de estudo. Nela, o leitor encontrará as seguintes contribuições:

No primeiro artigo, Anderson Fonseca analisa a ideia de uma colônia de formigas ser uma entidade consciente, a partir da semelhança do comportamento das formigas com o padrão de interação dos neurônios do cérebro humano. O autor fundamenta sua tese na definição de Thomas Nagel da consciência como *algo que é ser para um organismo* a fim de verificar se a colônia poderia ser um sujeito experiencial. Dentro dessa ótica, a colônia de formigas, se considerada um organismo, poderia ser um sujeito apto a ter experiências internas. Por isso, neste artigo, Fonseca desenvolve com base no Pampsiquismo,

critérios de análise da colônia, apresentando quais características ela deve atender para ser qualificada como consciente. Entretanto, o autor não defende que a colônia de formigas é um organismo genuíno, mas que, caso seja, terá algum tipo de experiência consciente. Assim, para tratar essa questão, Fonseca examina se a colônia é um indivíduo biológico, e, em seguida, se é possível que as mentes das formigas embasem a mente da colônia. Ao fim, ele conclui que, a partir do modo de interação das partes com o todo, pode haver mais de uma resolução para o problema, i.e., a consciência da colônia dependeria do modo como seus componentes estão física e fenomenalmente integrados.

Em seu artigo, Luke Roelofs também examina a possibilidade de colônias de formigas serem dotadas de mentes conscientes, mas a partir da perspectiva do combinacionismo pampsiquista, a tese de que mentes conscientes podem ser constituídas por outras mentes conscientes. Dentro dessa linha de pensamento, Roelofs defende tal conjectura como uma “premissa”, ou melhor, ele se propõe a investigar se colônias de formigas são conscientes, então *como são essas mentes?, como é ser uma colônia de formigas?* Assim, com o objetivo de desenvolver uma noção dos contornos fenomenais de uma colônia de formigas, o autor divide o capítulo em seis seções correspondentes a seis perguntas abarcando a ideia de a colônia de formigas ter um sistema nervoso distribuído, um corpo, uma mente, uma percepção, experiências fenomênicas, e certa relevância moral. No entanto, as seis questões acabam por fim se tornando uma *condição* para especular quais colônias podem ter mentes conscientes. Roelofs, então, mostra um quadro geral do que é a experiência de ser uma colônia de formigas e qual tipo de estrutura física fundamenta sua subjetividade.

O texto de Maria Luiza Iennaco e Daniel Friedman analisa como diferentes sistemas de medições empíricas e teorias científicas são

empregadas para investigar a consciência, entre as quais, a Teoria do Trabalho de Espaço Global e a Teoria da Informação Integrada. No entanto, os autores destacam o Princípio de Energia Livre e a Inferência Ativa como abordagens capazes de oferecer insights sobre a consciência quando aplicados, por exemplo, às colônias de formigas. A colônia de formigas é examinada, dentro dessa ótica, como modelo de uma perspectiva *bottom-up* da consciência, em que a relação entre comportamento coletivo e interação dinâmica com o ambiente são parâmetros para analisar a emergência da cognição em animais complexos. Em outras palavras, Iennaco e Friedman examinam como estruturas multidimensionais da consciência animal podem ser aplicadas a insetos eussociais, considerando colônias de formigas como um modelo empírico para complementar ideias teóricas de escalas de consciência *top-down*.

Em sua contribuição, Aryel C. Goes nos apresenta outra perspectiva sobre o tema central desta obra. Ele afirma que uma colônia de formigas teria uma consciência externalizada a partir da noção de que o conjunto das experiências das formigas individuais fundamentaria uma resposta higiênica coletiva contra a presença de patógenos a fim de manter a integridade do superorganismo. Isso significa que a colônia possui um sistema imunológico coletivo com uma memória interna das interações entre seus membros e corpos estranhos. Esse sistema agiria, a partir das experiências dos componentes, reconhecendo o que o constitui e excluindo o que ameaça a sua unidade. Mas o que seria então a mente consciente de tal superorganismo? Para Goes, ela corresponde ao conjunto das experiências das partes que sustentam a capacidade do grupo de responder como uma unidade às interferências em sua dinâmica interna. Por conseguinte, na visão do autor, essa consciência está distribuída nas formigas operárias que a compõem.

José Carlos Camillo, na mesma linha temática, apresenta uma discussão sobre a cientificidade de um fenômeno social particular: a memória coletiva. Recentemente, muitas investigações de diversas áreas têm se concentrado na temática da memória. Esse esforço coletivo é marcadamente interdisciplinar integrando Psicologia, Ciências Cognitivas e Filosofia. Uma parte desses esforços se direcionou ao tratamento da memória coletiva, fenômeno em que a organização e interação de pessoas em grupos pode impactar a retenção e retomada acurada de itens de testes de memória (listas de palavras, listas de desenhos, narrativas longas, etc.). O autor defende que o fenômeno da memória coletiva pode ser considerado um fenômeno científico, pois satisfaz os critérios de (i) ser um conjunto estável de propriedades causais, (ii) são replicáveis e (iii) suscetíveis de investigação por diversas metodologias. Além disso, estudos empíricos oferecem evidências para a legitimidade do fenômeno de memória coletiva ao mostrar sua autonomia com relação aos fenômenos mnemônicos individuais. Por fim, Camillo discute a extensão da afirmação de que a memória coletiva é um fenômeno científico. Mais especificamente, ele discute a legitimidade de uma inferência a partir da memória coletiva para a consciência coletiva. O autor argumenta negativamente a respeito dessa extensão. Ainda que memória dependa de consciência, o fenômeno da memória coletiva pode evidenciar que apenas existam as consciências dos indivíduos envolvidos e não, necessariamente, uma consciência coletiva. O fenômeno de uma consciência coletiva, ao menos ainda, não pode ser tratado como fenômeno científico, pois não satisfaz os critérios de (i) ser um conjunto estável de propriedades causais, (ii) ser replicável e (iii) ser suscetível de investigação por diversas metodologias.

No artigo de Valdenor Monteiro Brito Júnior é apresentada uma proposta de extensão da área da ontologia social. O autor destaca que as

propostas tradicionais de ontologia social, de teorias que buscam explicar a estrutura metafísica do mundo social, são significativamente ancoradas em pressupostos antropocêntricos. Mente e linguagem são os elementos centrais da construção da realidade social para essas teorias e os exemplos paradigmáticos de estados mentais e atos de fala são exclusivamente humanos, i.e., não são instanciados por nenhum outro ser vivo. Na tentativa de ampliar o quadro de entidades sociais legítimas que existem na realidade, Valdenor se utiliza de pesquisas da área da biologia para encontrar critérios mais abrangentes para serem aplicados nas teorias de ontologia social. A aposta se dá a partir do conceito de sociedade, enquanto um grupo de indivíduos pertencentes à mesma espécie e organizados de uma maneira cooperativa, e especificações sobre as causas e tipos de cooperação observadas em diversas espécies. A tese central do texto é baseada no comportamento de animais eussociais (formigas, abelhas, vespas, cupins), onde a estrutura social produzida por essas espécies não poderia ser descartada ou subestimada enquanto instância de entidade pertencente à ontologia social. O autor ainda discute uma série de objeções que buscariam descartar ou subestimar esses exemplos, argumentando que essas estratégias não são satisfatórias para demover a pertença desses casos ao conjunto de entidades sociais legítimas. Desta forma, deveríamos tratar igualmente as instâncias de entidades sociais produzidas por insetos eussociais e primatas humanos, pois ambos satisfazem o critério necessário para a existência de entidades sociais: a presença de certos tipos de comportamentos sujeitos aos princípios relativos à evolução da cooperação.

Luiz Paulo Cichoski, por sua vez, discute a tese da mente de grupos. O autor apresenta teorias em Filosofia da Mente que são compatíveis com uma posição realista sobre mente de grupos, ou seja, uma posição que considera legítimas as atribuições de estados mentais a grupos e

outras entidades sociais (como universidades, estados nacionais, famílias, tribunais de justiça, etc.). São apresentadas as teorias do funcionalismo, da mente estendida e da cognição socialmente distribuída e discute-se como cada uma delas poderia fundamentar uma posição realista neste debate.

Em “*Mentes de Grupos e Tipos Naturais*”, tradução do artigo de Robert Rupert, é apresentada uma posição crítica sobre mente de grupos. Rupert discute dois argumentos a favor da tese de que mentes de grupos são do mesmo tipo que as mentes individuais: um argumento via compromisso ontológico e outro baseado na similaridade do processamento cognitivo de indivíduos e grupos. Para o autor, ambos argumentos não são satisfatórios, pois não indicam como poderíamos identificar a igualdade de tipo entre os estados mentais de grupo e os estados mentais individuais. Propondo um critério para caracterização do que pode ser considerado cognição (um critério de integração cognitiva chamado de *probabilidade condicional de co-contribuição*), Rupert argumenta que as propostas realistas sobre mente de grupos não apresentam elementos que indiquem a satisfação deste critério. Portanto, não há igualdade entre os tipos naturais investigados pelas Ciências Cognitivas e as entidades discutidas pelos teóricos que defendem a existência de mentes de grupos.

Já em “*Introspecção em Mentes de Grupos, Desunião da Consciência e Pessoas Indiscretas*”, tradução do artigo de Eric Schwitzgebel e Sophie Nelson, se discute os limites da noção de introspecção. Essa noção é problematizada por meio da exploração de sua aplicação em contextos de mentes de grupos. Os autores utilizam uma estrutura ancilar, inspirada por uma arquitetura cognitiva sugerida pelo romance “*Justiça Ancilar*” de Ann Leckie. Esse sistema cognitivo é composto por um computador central e uma série de robôs conectados a esse computador,

fazendo com que a cognição dessa estrutura seja distribuída. Ao avaliar possibilidades de interpretação da relação entre os componentes desse sistema cognitivo distribuído, os autores problematizam a identificação de introspecção e comunicação em uma estrutura como essa. A dificuldade de determinação sobre qual tipo de relação se estabelece entre o computador central e os robôs, se de introspecção ou comunicação, acarreta em uma indeterminação no que diz respeito à fronteira da mente de um indivíduo. A indeterminação sobre a mente indica que entidades conscientes podem não ser discretas e contáveis, que a consciência possa ser gradativa e grupos concretos possam ter mentes indiscretas.

Por fim, antes de iniciarmos a discussão, gostaríamos de agradecer a todos os colegas que contribuíram com os textos que integram a obra. Um agradecimento especial ao professor João de Fernandes Teixeira por ter aceito escrever o prefácio. É uma enorme satisfação para nós ter a participação de uma das principais referências nacionais da Filosofia da Mente nesta publicação. Também gostaríamos de registrar e reconhecer a cuidadosa revisão realizada por Ana Luisa Lima Grein, cujo trabalho contribuiu significativamente para a qualidade final desta publicação. Agradecemos ainda a Eric Schwitzgebel, Sophie Nelson e Robert Rupert, bem como aos periódicos *Journal of Consciousness Studies* e *AVANT*, pela autorização para a publicação das duas traduções. Destacamos, ainda, o apoio financeiro do Programa de Pós-Graduação em Filosofia da Universidade Federal de Mato Grosso, por meio do Programa de Desenvolvimento da Pós-Graduação (processo nº 88881.808937/2023-01), que tornou possível a concretização desta obra.

1

É A COLÔNIA DE FORMIGAS UM ORGANISMO CONSCIENTE? ¹

Anderson Luiz do Vale Fonseca

Neste artigo, eu pretendo, examinar teoricamente se uma colônia de organismos invertebrados possui experiência consciente como uma unidade. No caso, em específico, uma colônia de formigas. Para isso, me embaso na ideia de consciência como “algo que é sentir-se para o organismo” (Nagel, 1979, p. 166)². Esse conceito, se aplicado à colônia de formigas, sugere ser possível ela manifestar aspectos subjetivos da consciência. No entanto, conforme os critérios ontológicos da experiência, estabelecidos pelo filósofo americano Thomas Nagel (Ibid., p. 166), a colônia de formigas precisa ser primeiro um organismo, para em seguida ter um sentimento de si como um organismo. Sendo assim, uma colônia de formigas se for um organismo genuíno teria uma consciência fundamentada no conjunto de suas partes conscientes. Isto requer, evidentemente, uma análise do conceito de indivíduo biológico e sua relação com a experiência subjetiva de um sistema vivo. No caso, seria a colônia de formigas um organismo genuíno? Seria ela, pois, um sujeito experiencial? Em outras palavras, há algo que é ser como uma colônia de formigas?

A minha análise se embasa no Pampsiquismo, a tese de que os constituintes físicos básicos do mundo natural têm propriedades mentais. De modo específico, em uma versão desse argumento,

¹ Artigo originalmente publicado em Griot : Revista de Filosofia, [S. l.], v. 23, n. 1, p. 70–86, 2023. DOI: 10.31977/grifi.v23i1.3220. Disponível em: <https://periodicos.ufrb.edu.br/index.php/griot/article/view/3220>

² Todas as citações neste texto foram traduzidas do original pelo autor.

conhecido como Pampsiquismo Combinacionista, segundo o qual “as propriedades experienciais de um sujeito consciente são às vezes meras combinações das propriedades experienciais de outros sujeitos que o compõem” (Roelofs, 2019, p. 6). A consciência animal, sob esta ótica, poderia ser uma subjetividade composta por componentes individualmente conscientes. É possível então afirmar que sujeitos conscientes podem compor sujeitos conscientes adicionais.

Deste modo, uma Colônia de formigas poderia ser um sujeito consciente. Ou seja, as experiências fenomenais de suas partes se combinariam para constituir a experiência fenomenal dela como uma entidade viva. Contudo, como já afirmei, é importante analisar se a colônia de formigas é um indivíduo. Em seguida, verificar se as formigas que constituem a colônia são sujeitos conscientes. A partir daí, examinar como combinam suas experiências para fundamentar a experiência da colônia.

Não estou aqui, de antemão, afirmando que a colônia tenha uma consciência, nem é meu objetivo seguir uma linha de pensamento que leve a esse fim. O que estou abordando é que, se a colônia for um organismo, a tese pampsiquista se aplicará a ela também. Isso envolve, portanto, a problematização dos conceitos de Composição material e o conceito de Consciência, i.e., a conexão entre a forma física do ser vivo e a subjetividade.

Para o Pampsiquismo, desde que as partes físicas se combinem de modo adequado temos um indivíduo. Na esteira deste raciocínio, um organismo consciente seria um sistema material organizado de *forma correta*, i.e., de modo que as partes estejam conectadas o suficiente para constituir um todo unificado. Entretanto, a Composição admite que nem todo sistema vivo estaria organizado da *forma correta* para ser um sujeito de experiência. Ou seja, um organismo desprovido de uma rede

nervosa centralizada não estaria apto a sentir o meio internamente. Mas até esta ideia parece controversa. Ser um organismo é bastante para ter uma experiência subjetiva? Se a resposta for positiva, então a ideia de associar exclusivamente a consciência a sistemas nervosos pode não ser exata. Nesse sentido, uma célula seria um sujeito experiencial? O Pampsiquismo, encontra, portanto, o seguinte obstáculo: para definir o que é a mente, precisamos entender o que é o corpo. E, se não há um conceito preciso de indivíduo biológico, tão pouco haverá um de consciência. Nesse caso, qualquer forma de organização material concebida como sistema vivo estará adequada para ser vista como consciente em algum grau.

O Problema da Composição envolve, portanto, a inconsistência do conceito de organismo. Em linhas gerais, se a noção de indivíduo é vaga, por sua vez, a de consciência também será. Ademais, Nagel não apresenta um princípio de organização geral que esclareça a combinação correta das partes para formar uma entidade subjetiva. Por conseguinte, qualquer relação entre parte e todo que atenda ao conceito de organismo de uma dada teoria científica, pode ser qualificada como um indivíduo genuíno. Contudo, isso significa que talvez não seja necessário ter um cérebro para ser um sujeito consciente, mas o tipo de conexão entre o sistema e seus subsistemas bastaria, como uma rede de comunicação eficaz. Sendo assim, uma colônia de organismos unicelulares como a *Caravela-Portuguesa*, cujo corpo é similar à de uma água-viva, teria alguma experiência como unidade. Do mesmo modo, uma colônia de insetos sociais. Estou, nesse caso, me baseando em sistemas que não têm uma rede neural biológica conectando suas partes a um núcleo central.

Soma-se a isso que, o Pampsiquismo defende uma perspectiva irrestrita da composição. Segundo esse ponto de vista, os componentes

físicos que constituem um indivíduo, podem ser recombinados para compor outro indivíduo. Entretanto, a noção de parte e de todo também é vaga. Os átomos que compõem uma pessoa humana são *parte* dela, e, por sua vez, ela é *parte* de uma nação ou comunidade. Ora, se a composição irrestrita for verdadeira para qualquer relação parte-todo, um indivíduo pode ser composto por outros indivíduos, desde que exista uma interação capaz de produzir uma unidade. Por conseguinte, o Pampsiquismo parece aceitar a ideia de que um sistema vivo possa compor outro. O importante é que os componentes estejam ordenadamente arranjados na composição do indivíduo adicional (Nagel, 2004, p. 43-44; Roelofs, 2019, p. 21). Mas como determinar se um organismo está estruturado de modo apropriado para ter uma consciência? A água-viva que não tem um sistema nervoso centralizado, mas distribuído nos tentáculos, teria algum tipo de consciência? A formiga, ou a abelha, mesmo sem um córtex neural superior, tem experiência subjetiva? Uma colônia, seja ela feita de corpos unicelulares ou multicelulares, teria consciência como um indivíduo?

1. As quatro premissas do Pampsiquismo

De acordo com o filósofo Thomas Nagel, em seu ensaio *Panpsychism*, publicado no livro *Mortal Questions*, o Pampsiquismo³ tem como fundamento quatro premissas simples (Nagel, 1979, p. 181-182; Cleve, 1990, p. 215), a saber:

³ Em linhas gerais, “é a tese de que algumas entidades microfísicas são conscientes, isto é, há algo que é como ser um quark ou um fóton ou um membro de algum outro tipo físico fundamental” CHALMERS, David. *Panpsychism and Panprotopsychism*. 2013, p. 1.

1. *Composição material*, ou, *Antidualismo*, segundo a qual qualquer organismo vivo é um sistema material complexo, consistindo em um grande número de partículas combinadas de modo especial;
2. *Antirreduccionismo*, no qual assevera que os estados mentais não são propriedades físicas do organismo e nem inteligivelmente derivadas apenas de suas propriedades físicas;
3. *Realismo* ou *Antieliminativismo*, ou a visão de que estados mentais são propriedades genuínas do organismo e não são propriedades de nada em absoluto;
4. *Antiemergentismo*, ou o entendimento da não existência de propriedades verdadeiramente emergentes de sistemas complexos, mas que são inteligivelmente derivadas das propriedades de seus constituintes e de seus efeitos um sobre os outros quando combinados.

Neste sentido, o Pampsiquismo nageliano (iremos chamá-lo assim), afirma que alguns constituintes físicos devem ter propriedades mentais não-físicas das quais quando combinadas em uma estrutura correta fundamentam a emergência de uma mente consciente. A experiência subjetiva, contudo, não seria uma propriedade emergente no sentido forte do termo, ou melhor, como algo que surge acima da estrutura física, e sim, uma característica intrínseca das entidades físicas básicas. Contudo, quando essas propriedades são reunidas aparece a consciência da entidade biológica. Por isto, Nagel diz que a mente é uma propriedade biológica, pois somente a *persona* resultante teria uma consciência genuína. Não haveria, portanto, a consciência das partes e a consciência do todo. De acordo com Nagel (2004, p. 48), só há a consciência do organismo e nada mais. Se somente entidades biológicas têm estados mentais conscientes, então o princípio realista

implica na existência de uma conexão necessária entre a neurofisiologia do sistema vivo e a aparência dos estados mentais. Contudo, para que o vínculo seja necessário, as propriedades mentais devem ser não-emergentes, ou seja, as propriedades mentais do sistema vivo devem estar fundadas nas propriedades micromentais das entidades microfísicas que o compõem. De modo análogo, as propriedades mentais da Colônia como uma unidade biológica estaria embasada nas propriedades mentais das formigas.

Assim, para os propósitos do presente estudo, considero que as quatro premissas do Pampsiquismo podem ser interpretadas como condições para que uma unidade tenha experiência consciente. Em síntese: uma determinada ordem material complexa poderá ser consciente se as entidades físicas que a compõem também forem conscientes. No entanto, precisam estar dispostas de modo a constituir um organismo. Por isso, essas quatro premissas se tomadas como critérios, podem ser aplicadas à Colônia de formigas. Se a colônia atender a estes parâmetros terá algum tipo de experiência subjetiva. Portanto, ela precisa, segundo a (o):

- a. *Composição Material: Ser uma estrutura cuja ordenamento das partes é suficiente para ter uma consciência.* Ou seja, a Colônia de formigas precisa ser realmente um organismo.
- b. *Antirreduccionismo: Uma explicação da Mente da Colônia com base apenas na descrição do comportamento de suas partes não é suficiente.*
- c. *Realismo: Se a Colônia é um organismo genuíno, segue que ela tenha estados mentais conscientes.*
- d. *Antiemergência: A Mente da Colônia não é uma propriedade emergente, mas deriva diretamente da soma das mentes individuais*

das formigas. Por conseguinte, as formigas têm que ser conscientes.

Em suma, a colônia sendo um organismo, terá uma consciência e mesmo que sua subjetividade derive diretamente da soma das subjetividades individuais, não é redutível a uma explicação unicamente comportamentalista. Ainda que a emergência da colônia seja elucidada a partir do comportamento cooperativo das partes, não é suficiente. É importante, acrescentar a possibilidade de os componentes terem estados subjetivos de consciência. O ponto de partida de análise da presente pesquisa baseado na teoria pampsiquista, sustenta que micromentes podem compor macromentes, quando combinadas em uma estrutura integrada.

No entanto, afirmá-las requer que assumamos dois princípios: o de aninhamento e o de combinação. De acordo com o Princípio de aninhamento, a consciência contém como parte de si, outros sistemas conscientes, ou pode estar contida em outros sistemas conscientes. De acordo com o Princípio da Combinação, a consciência é composta de partes conscientes reunidas de modo inteligível. Neste âmbito, a consciência de uma Colônia de formigas estaria embasada no conjunto das consciências individuais que a compõem.

Na esteira deste raciocínio, para que o argumento combinacionista faça algum sentido, acho razoável que, primeiro a colônia seja um organismo genuíno e, em segundo, que as formigas sejam sujeitos experienciais. Deste modo, sua mente seria estruturalmente aninhada. Buscarei, então, examinar cada um destes três pontos.

2. A colônia de formigas é um organismo?

Para responder a esta pergunta, considero a Teoria da Organismalidade, defendida por David C. Queller e Joan E. Strassman (2009), cuja pesquisa em biologia evolutiva se tornou crucial por fornecer uma agenda empírica para entender as razões da individualidade biológica, fornece bases para a análise. Sua ideia central é de que um organismo é produto de uma alta cooperação associada a um forte controle dos conflitos intraorganismais (Queller & Strassman, 2009, p. 3144; Strassman & Queller, 2010, p. 605). Nesse sentido, colônias de insetos sociais ou comunidades multiespécies com alta cooperação seriam integrados o suficiente para constituírem organismos genuínos.

Essa teoria dá continuidade ao pensamento do mirmecólogo norte-americano William Morton Wheeler (1911), o qual estabelece como argumento para a organismalidade da colônia a sociogenia, a ideia de que todo organismo manifesta uma forte predileção por buscar outros organismos e assimilá-los ou cooperar com eles para compor um indivíduo biológico mais abrangente e eficiente, a fim de garantir a própria sobrevivência. Na visão de Wheeler (1911, p. 307-308), a Colônia de formigas é um organismo genuíno.

Para ele a colônia de formigas seria um organismo no sentido de um indivíduo fisiológico, definido por suas propriedades estruturais – como ter partes heterogêneas e especializadas – bem como pelas propriedades funcionais – como a capacidade de desenvolvimento, reprodução e autorreparação. Isso significa que a colônia como unidade é caracterizada pela separação definitiva das operárias, desde a fase embrionária, em estéreis que desempenharão papel idêntico ao das células somáticas e em reprodutivas, as quais exercem a mesma função

que as células germinativas no sistema biológico. Sob este ponto de vista, Wheeler escreve (1911, p. 308):

Um organismo é um sistema complexo, absolutamente coordenado e, portanto, individualizado, de atividades dirigidas principalmente a obter e assimilar substâncias do ambiente, a produzir outros sistemas similares, conhecidos como descendentes, e a autoprotoger o sistema e, geralmente, também seus descendentes de perturbações que emanam do ambiente.

Esse conceito implica que uma entidade autônoma é caracterizada pela coordenação e integração de seus componentes, os quais produzem este indivíduo destacado do meio. A Teoria da Organismalidade segue a mesma linha de raciocínio de Wheeler, ao afirmar que seres vivos são formados por diferentes níveis de grupos sociais que cooperam entre si, enfatiza a adaptação como aspecto central do ser vivo. Ou seja, cooperação elevada e conflito controlado ou ausente entre as partes é dizer que o organismo é o objetivo da adaptação. Portanto, um sistema vivo é uma unidade de adaptação. É possível, então, que exista grupos com alta cooperação e alto conflito, constituindo, pois, uma sociedade; enquanto, grupos com baixa cooperação e baixo conflito, apenas formam grupos simples (Strassman & Queller, 2010, p. 605-607). Porém, como dizer quais unidades são organismos verdadeiros?

Queller e Strassman (2009, p. 3144; 2010, p. 608) afirmam que uma distinção adequada se baseia na característica das partes, se elas são indivíduos semelhantes ou diferentes. Como resultado, tem-se dois tipos de organismos: fraternos e igualitários. Os organismos fraternos são aqueles compostos por múltiplos indivíduos de uma mesma espécie ou por indivíduos clonais, que se agrupam com o objetivo de transferir o *pool* gênico à geração seguinte. De maneira oposta, os organismos igualitários são aqueles compostos por indivíduos biológicos diferentes que estabelecem uma relação mutualista. Entretanto, esse modo de

agrupamento pode ocorrer em vários níveis, ou seja, uma entidade biológica pode ser uma reunião de indivíduos fraternos e igualitários, a exemplo do homem que é constituído por células e outros micro-organismos.

O conceito de organismos fraternos toma como paradigma organismos multicelulares, como gatos e cães. Em linhas gerais, os organismos paradigmáticos fraternos são organismos clonais, em que cada parte é geneticamente idêntica a todas as outras. Nesse contexto, a alta cooperação é produto do estreitamento parental. Consequentemente, temos uma unidade fisiologicamente coesa, em que todas as células que compõe a unidade colaboram para garantir a reprodução das células germinativas. Entretanto, em um organismo multicelular também há conflitos, é possível mutações em células somáticas que as desconectam do grupo convertendo-as em cânceres. Assim, uma unidade biológica não apenas é efeito da alta cooperação como também do conflito entre componentes, embora, os conflitos entre as partes, ou são de menor dimensão, ou controlado.

Um exemplo de organismo fraternal é o *D. discoideum*, uma ameba que, dependendo das condições do ambiente, pode migrar de uma vida solitária para uma social, ao se agregar a outras próximas engendrando, assim, um organismo multicelular. Imaginando que durante o período multicelular haja níveis de cooperação assim como níveis latentes de conflito, porém se o conflito é baixo ou esporádico, segue-se então que *D. discoideum* é um organismo verdadeiro (Queller & Strassman, 2009, p. 3145; 2010, p. 608; 2011, p. 597-607). Dependendo das condições ambientais, ela pode sair de um modo de vida solitário para o gregário/social.

Além dele, colônias de indivíduos multicelulares como os sifonóforos cnidários, de modo específico, a *Physalia physalis*, também podem ser considerados organismos, dado que o baixo conflito e a alta

cooperação são resultados da clonalidade de um único indivíduo unicelular. Por conseguinte, os grupos clonais se especializam funcional e morfológicamente em quatro tipos diferentes de pólipos: a) pneumatóforo, uma vesícula cheia de ar para garantir a flutuação; b) gastrozoides, que fazem o papel de aparelho digestivo da colônia; c) dactilozoides, que funcionam como tentáculos; d) e os gonozoides, destinados à reprodução. Estes organismos são partes dependentes entre si, além de morfológicamente estarem associados, de modo que se torna impossível viverem separados. No entanto, a *P. physalis*, na ótica tradicional, não é conceituada como uma entidade biológica, e sim um agregado de pólipos. O argumento de Queller e Strassman (2009, p. 31247; 2010, p. 609), em oposição a esse ponto de vista, afirma que o sifonóforo é genuinamente um organismo, dada a elevada integração das partes.

Em contrapartida, grupos cooperativos não-clonais podem exibir mais cooperatividade e controle de conflitos do que grupos clonais. Uma colônia de insetos sociais, por exemplo, em razão do grau de parentesco entre irmãs, pode ser qualificada como organismo ou não. Considere a hipótese da haplodiploidia, que sustenta a ideia de que em uma colônia de formigas, o coeficiente de relação genética entre as irmãs é de 0,75 enquanto elas em relação aos seus irmãos é de 0,25. Essa proximidade genética se torna essencial para a cooperação entre as formigas operárias (fêmeas), tendo em vista que os machos pouco colaboram. Todavia, como apontado, nem todas as colônias são organismos. À vista disso, Strassman e Queller (2010, p. 609) estabelecem quatro critérios para analisá-las, se um certo sistema é um organismo:

1. Quando o conflito é forte o suficiente, colônias não são organismos.
2. Quando o conflito é moderado, são organismos.

3. Quando o conflito é zero e a cooperação é igualmente zero, não são organismos.
4. Quando o conflito é baixo e a cooperação é alta, são organismos.

Nesse contexto, se um grupo de formigas se torna antissocial, a colônia perde a organismalidade. O que importa à organismalidade é a cooperação, ou sociabilidade, entre os componentes. Se houver conflito, deverá ser em pequenas proporções, ou localmente, e se não, no mínimo controlável. Assim, estes aspectos são a medida para determinar quais agrupamentos sociais constituem entidades biológicas.

O conceito de organismos igualitários, por sua vez, toma como paradigma a parceria da célula eucariótica, composta pela mitocôndria e a célula hospedeira. As células eucarióticas são a base para a variabilidade de formas vivas. A partir desse modelo, um organismo pode também ser algo constituído de diferentes espécies. Por essa linha de raciocínio, a cooperação sexual entre parceiros de diferentes espécies é o bastante para defini-los como organismo. O caso mais ilustre é o Peixe-diabo (*Caulophryne jordani*), em que o macho (com cerca de 1,5 cm) é “parasita” da fêmea (com aproximadamente 18 cm). O macho ao se fixar em definitivo no corpo da fêmea, atrofia todos os órgãos exceto as gônadas (onde se produz as células sexuais), alcançando assim o estágio “parasitário”. Existe, pois, uma relação mutualista: de um lado, o macho se alimenta do sangue da fêmea; do outro, fornece seus gametas para a fecundação dos óvulos. Essa parceria demonstra, portanto, que os organismos podem ser produzidos pela integração de indivíduos não relacionados (Strassman & Queller, 2010, p. 610).

Por conseguinte, tanto os organismos fraternais quanto os igualitários esboçam três tipos de condições para a organismalidade em grupos sociais (Strassman & Queller, 2010, p. 611):

Em primeiro lugar, deve haver vantagens sinérgicas para a cooperação que a tornam mais do que um jogo de soma zero. Em segundo lugar, algumas condições devem alinhar largamente os interesses reprodutivos das partes. Finalmente, quando os conflitos potenciais permanecem, eles devem ser suprimidos por algum mecanismo para que a organismalidade seja alcançada.

Essas condições são visíveis na colônia de insetos sociais dado que o altruísmo fornece os benefícios evolutivos para o grupo. A cooperação só é ameaçada em situações como a morte da rainha. Em casos assim, se não houver operárias que possam assumir o lugar da rainha, a colônia também morrerá. Ademais, é possível afirmar que o feromônio de agregação liberado pela rainha seja um meio de reduzir o conflito e manter a cooperação.

Em suma, um organismo é uma unidade social definida pela elevada cooperação das partes. Esse argumento pode ser estendido aos demais animais coloniais produzidos por grupos sociais clonais e não-clonais, assim como a outros animais incluindo o homem. Nesta perspectiva, tendo em conta o nível de cooperação e conflito, haveria a possibilidade de agrupamentos que se aproximam da unidade, do mesmo modo a existência de agrupamentos que alcançam a unidade (Strassman & Queller, 2010, p. 614). Neste sentido, ainda que a Colônia de formigas apresente características que atendam as condições teóricas para ser um organismo genuíno, tendo em vista o fato de as formigas não estarem estruturalmente conectadas como ocorre com o *P. physalis*, na formação de um corpo multicelular, mas espalhadas pelo ambiente, favorece a visão da colônia como se encontrando entre a agregação e a individualidade. Deste modo, a colônia de formigas seria um organismo indeterminado.

3. As formigas são conscientes?

Supondo, então, que as formigas são conscientes, em que substrato neural ela está fundada? Em que local do cérebro as informações sensoriais processadas são integradas na experiência consciente? E, qual a relação entre a experiência da formiga e a organização colonial? Uma abordagem teórica que defende a existência de consciência em animais invertebrados é o argumento de Colin Klein e Andrew B. Barron (2016), cuja importância para o estudo da mente está em apresentar uma nova maneira de entender as origens neurobiológicas da experiência subjetiva. De acordo com estes pesquisadores, a consciência dos insetos ocorre em uma região sub-cortical do cérebro, o mesencéfalo. No cérebro mediano as informações sensoriais interoceptivas (estímulos provenientes do corpo) e exteroceptivas (estímulos externos ao corpo) são combinadas em um modelo neural do corpo em movimento no ambiente (Klein & Barron, 2016, p. 3). Sendo assim, os insetos teriam um sistema neural integrado capaz de suportar uma experiência consciente (Merker, 2007, p. 70-74).

Duas características do mesencéfalo são fundamentais para sustentar a experiência subjetiva em insetos. Na primeira, o processamento integrado de informações espaciais permite ao animal em movimento eliminar a ambiguidade entre a saída motora e a entrada sensorial causada por sua ação no ambiente. As diversas informações sensoriais incluindo posição, orientação e movimento do corpo (ou somato-espacial) são unificadas em um modelo neural permitindo ao animal mover-se de maneira organizada no espaço e reagir às mudanças locais. A segunda característica é a integração de informações sobre as necessidades fisiológicas do animal com a localização e disponibilidade de recursos resultar na seleção de ações

eficientes a serem executadas pelo animal. Enfim, o mesencéfalo embasa a tomada de decisão do indivíduo e o planejamento de ações (Klein & Barron, 2016, p. 3-8).

Seguindo essa linha de raciocínio, Klein e Barron (2016, p. 5) argumentam que a representação unificada e centrada do mundo na perspectiva do inseto é suficiente para que ele tenha experiência consciente. Ou seja, esses animais são sencientes, são aptos a sentir (*feeling*) internamente sua localização no espaço e responder aos estímulos externos (Klein & Barron, 2016, p. 10-11; Søvik & Perry, 2016, p. 2). Ademais, tendo em conta que a consciência envolve a faculdade de perceber algo, ou, depende da existência de objetos externos, o fato de os insetos terem circuitos de atenção no protocérebro sugere que se tornam conscientes do ambiente em torno (Feinberg & Mallat, 2016, p. 4).

É possível, a partir disso, que se interprete o argumento de Klein e Barron (2016) da consciência em insetos como restrito a um senso rudimentar do espaço. Na verdade, para que fique ainda mais claro, não se trata apenas de se mover no espaço e responder aos estímulos ambientais, e sim, a combinação da informação somatoespacial em uma representação interna (Søvik & Perry, 2016, p. 2-3). Considerando então que a consciência é apoiada por estruturas subcorticais, o número de neurônios por mais relevante que seja, não faria muita diferença sem uma organização funcional apropriada

Mas Klein e Barron (2016, p. 10-11) não sugerem que os insetos sejam capazes de monitorar seus próprios processos cognitivos, nem que sejam autoconscientes. Eles não teriam o mesmo tipo de experiência subjetiva que os mamíferos com córtex cerebral têm. Na verdade, teriam uma consciência primária, um nível básico de experiência (Søvik & Perry, 2016, p. 4; Feinberg & Mallat, 2016, p. 4). A questão envolvendo o conceito de “nível” é a imprecisão quanto ao número de neurônios suficiente para se

atribuir consciência. Possivelmente, a ocorrência da experiência subjetiva não depende da quantidade de células nervosas, mas sim do modo como o sistema está organizado.

Por fim, se existe a possibilidade de as formigas serem capazes de ter um aspecto subjetivo da experiência a partir da integração das informações sensoriais químicas, visuais e motoras em uma representação interna, então a consciência seria uma explicação extra não somente para a tomada de decisão, mas também para a organização colonial. Como resultado, a experiência consciente forneceria outra percepção do comportamento das formigas (Klein & Barron, 2016, p. 7). Agora, falta dizer de que modo seria integrada a consciência das partes para constituir a experiência unificada do todo

4. É possível reunir mentes?

Para fundamentar a ideia de soma de consciências individuais em uma consciência conjunta, proponho a seguinte experiência de pensamento: Suponha agora que a ameba *D. discoideum* tem experiência consciente. Essa ameba unicelular passa por quatro estágios em sua curta vida: vegetação (fase unicelular), agregação (constituição da colônia ou do organismo multicelular), migração (o corpo se transporta para uma região com mais nutrientes disponíveis) e culminação (desagregação)⁴. Sabendo disso, vamos imaginar que cada estágio corresponda a um tipo de consciência resultante do nível de integração física e experiencial.

⁴ Aqui tomei como base a tese de doutorado de Daniel Schardosim Calovi, onde o autor apresenta uma análise criteriosa dos estágios do ciclo de vida social da ameba *D. discoideum* simulados computacionalmente. Ver CALOVI, Daniel Schardosim. *Simulação do ciclo de vida social da ameba Dictyostelium discoideum*. 82 f. Tese (Doutorado em Ciências) – Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011. Disponível em: <https://lume.ufrgs.br/handle/10183/31612>. Acesso em 20 de agosto de 2021.

No início, a consciência ainda é individual, abrangendo as experiências do organismo unicelular como as sensações de detecção de nutrientes no ambiente e o prazer de se alimentar dos micro-organismos disponíveis. Quando, porém, a quantidade de nutrientes se torna escassa, cada indivíduo emite um sinal químico convocando os demais a agregar-se, produzindo assim uma colônia. Nesse estágio inicial cada organismo está cooperando com o outro em um nível fraco de integração física (isto é, ainda não estão anatomicamente conectados) e fenomenal (isto é, a experiência partilhada é parcial). Como resultado, se tem uma espécie de consciência coletiva caracterizada por estar entre a consciência do indivíduo unicelular e a do grupo. No momento em que se alcança uma forte integração entre as partes é gerado um organismo multicelular, ao mesmo tempo a consciência da unidade. Aqui, as experiências das partes estão combinadas para compor a experiência do organismo multicelular. E isso dura até que se encontrar um local com um volume de nutrientes suficiente para cada ameba, causando, a seguir, a desagregação.

Supondo que ideia de uma mente feita da combinação de outras mentes seja verdadeira, ela exige que o limite entre a experiência dos componentes e a experiência da unidade sejam conceitualmente bem definidas. No exemplo do *D. discoideum*, em que a agregação dos organismos unicelulares resulta na constituição de um organismo multicelular, se é verdade que as partes da colônia são conscientes, não se segue que a unidade biológica tenha consciência. É de pressupor que a colônia se resume a uma agregação de indivíduos por oferecer vantagens seletivas, sem disso resultar em uma mente unificada. Cada membro pode ter um caráter qualitativo da experiência e jamais compartilhá-lo com os outros. De modo algum haveria uma junção das experiências individuais para dar existência à experiência da entidade biológica.

Aqui estou supondo a seguinte ideia do pampsiquismo constitutivo: sujeitos experienciais, podem compor outros sujeitos experienciais (não levando em conta detalhes como ter um sistema nervoso) (Goff, 2017, p. 297). Neste caso, o *D. discoideum* nos leva ao chamado “problema da combinação”. De que maneira micro-consciências se combinam para constituir macroconsciências? Este problema se trata de entender a unidade da consciência em uma estrutura material composta (o cérebro), de como as informações sensoriais processadas nos conjuntos de neurônios são unificadas na experiência consciente do sujeito.

Para ilustrar essa questão, vejamos o experimento do cérebro dividido (*split-brain*). Em síntese, um paciente com epilepsia passa por uma cirurgia chamada calossotomia, a qual consiste no corte do corpo caloso, cuja função é conectar os dois hemisférios cerebrais direito e esquerdo permitindo a eles que se comuniquem entre si. Após a cirurgia temos um sujeito com os dois hemisférios desconectados. O evento é curiosamente interessante, visto que o lado direito e o lado esquerdo do cérebro estão separados surge um único indivíduo com uma consciência desunificada (Goff, 2019, p. 148-154; Roelofs, 2019, p. 207).

Nagel (1979, p. 159-164), por exemplo, afirma que a experiência unificada da consciência possivelmente é resultado da integração de duas mentes diferentes. Fundamentado no experimento do cérebro dividido, o filósofo chega à ideia de a consciência do sujeito ser efeito da combinação das consciências das entidades físicas que o constitui. Se cada hemisfério cerebral tem uma mente, parece racional supor que as operações mentais de cada lado podem ser também separadas formando mentes distintas.

A questão que surge no Pampsiquismo, de acordo com Goff (2019, p.148), é: “Como se passa de uma consciência desunificada (a consciência

das partes isoladas) para uma estrutura neural com consciência unificada?” Possivelmente, pela forma como os constituintes básicos foram arranjados (Nagel, 2004, p. 48). Esta forma adequada, conforme a premissa da composição, seria o organismo. Mas a possibilidade de organismos unicelulares, na hipótese de terem consciência, ao constituir um organismo multicelular produzirem a consciência desse indivíduo depende de dois fatores: do quão estreita é a relação entre os componentes e da experiência das partes ser passível de compartilhamento a um certo nível entre elas e não totalmente privativa (Roelofs, 2019, p. 63-64; Goff, 2017, p. 292-293; Roelofs & Goff, 2020, p. 7). Caso contrário, haverá a consciência isolada dos grupos de sujeitos sem um sujeito adicional consciente.

Em sua obra *Combining Minds*, o filósofo Luke Roelofs (2019, p. 20), defende o combinacionismo, a tese de que “as propriedades experienciais do todo podem, em alguns casos, serem fundadas e explicadas pelas de suas partes”. A consciência humana, sob esta ótica, poderia ser uma subjetividade composta por componentes individualmente conscientes. Uma colônia de insetos himenópteros sociais, em caso positivo, teria sua consciência fundada na consciência de seus membros? Seria o “eu” da colônia o mesmo das partes ou algo diferente? Ou não seria ela uma entidade subjetiva? A resposta deve se apoiar na seguinte conceituação do combinacionismo:

As propriedades experienciais de um sujeito consciente são às vezes meras combinações das propriedades experienciais de outros sujeitos que o compõem. Um modo mais intuitivo de colocar a definição de “combinação” é dizer que uma característica minha é uma combinação das características de minhas partes se essas outras características, e a forma como estão conectadas, ambas são suficientes para fundamentar o fato de eu estar tendo a característica em questão e podem ser apeladas em uma explicação do porquê eu tê-las (Roelofs, 2019, p. 6).

A princípio, afirmar que as propriedades experienciais de um sistema estão fundadas nas propriedades experienciais de seus constituintes significa dizer que uma vez tendo *B* isso é suficiente para termos *A*, ou melhor, as propriedades de um estão fundadas nas propriedades do outro. Nessa perspectiva, a consciência de *A* está ligada à consciência de *B*, de modo que sem *B* não existe *A* (Roelofs, 2019, p. 23).

Em segundo ponto, ao dizer que as propriedades do todo são explicadas pelas propriedades das partes significa que ao explicar *B* temos uma compreensão completa de *A*. Ou seja, a análise dos componentes em um determinado momento é suficiente para explicar as propriedades do sistema naquele exato momento. Nesse âmbito, Roelofs (2019, p. 23) argumenta que “a consciência é uma mera combinação das partes se e somente se é tanto fundamentado nelas e suas interrelações como também explicado por elas e suas interrelações”.

Mas o que são as experiências das partes? O que é a experiência do todo? Ora, retomando mais uma vez o experimento mental do *D. discoideum*, a experiência dos constituintes pode ser entendida como uma experiência desunificada (no estágio vegetal), algo que é como ser cada indivíduo, enquanto a experiência do todo, evidentemente, corresponde a uma experiência unificada (no estágio de agregação), *algo que é ser como essa entidade* (Nagel, 1979, p. 164, 182; 2004, p. 81). Logo, temos que supor ser o padrão organizacional instanciado pela cooperação das partes, a estrutura em que aparece a consciência.

Em geral, a análise da consciência é feita pela ótica dela como unidade genuína, de tal modo que não é possível vê-la como um conjunto de partes. No entanto, o argumento que aqui sustento, tendo como base o pampsiquismo, é: o organismo e a consciência são unidades compostas por constituintes relativamente integrados. Nessa linha, eu

sustento a visão do combinacionismo, o qual pode ser definido em termos de uma relação parte-todo, em que as propriedades experienciais das partes, quando combinadas, fundamentam e explicam as propriedades experienciais do sujeito consciente (Roelofs, 2019, p. 30).

Tal doutrina se opõe ao argumento da Mente Chinesa, proposto pelo filósofo Ned Block (2002, p. 96-97), que afirma não ser coerente a ideia de uma subjetividade composta. Suponha, pois, que o governo da China convença seus cidadãos de simularem um cérebro humano usando rádios transmissores. Cada cidadão recebe um dispositivo para enviar e receber sinais, além de instruções muito específicas como “ao receber o sinal I, do dispositivo S_i, emita um sinal O para o dispositivo S_i” (Block, 2002, p. 97). Suponha ainda que as entradas e saídas dos sinais ocorram simultaneamente, compartilhadas com centenas de milhares de indivíduos, de modo similar ao sistema neural humano. Em seguida um androide é conectado aos aparelhos a fim de ser controlado pelos cidadãos de maneira igual às entradas sensoriais e saídas motoras no cérebro humano. Consequentemente, os sinais enviados por milhões de cidadãos acabam por produzir uma máquina com comportamento inteligente. A questão que isto suscita é: o conjunto destes cidadãos está consciente como sendo um único indivíduo?

Para Block (2002, p. 97), é contraintuitivo afirmar que tenham uma experiência consciente em grupo. É absurdo. Não existe, neste caso, uma integração das experiências individuais. O que há, é um *gap* (hiato) entre a consciência das partes e o todo. Ademais, uma entidade consciente não pode conter outra entidade consciente como parte ou estar contida em. Nesse âmbito, ou a consciência existirá no todo, ou existirá nas partes, jamais em ambas. Mas não é o que pensa Roelofs. Segundo ele (2019, p. 43), a impossibilidade de emergir uma mente

unificada deve-se à falta de uma estrutura específica que integre as informações produzidas pelo agregado dos componentes.

Goff aborda essa questão, porque se a consciência unificada é resultante da combinação de outras consciências, é racional, então, pensar que existe entre os componentes uma espécie de ligação fenomenal (*phenomenal bonding*), isto é, uma relação entre campos fenomenais (ou experiências individuadas) que dão origem às estruturas conscientes a partir de uma experiência composta. A questão do que ele chama de Composição Fenomenal é definida pelo modo como os sujeitos com fenomenalidade se conectam para constituir outros sujeitos, exige que se determine a *forma adequada* em que a composição é realizada, além de quais partes mantém essa relação fenomenal. Posto isto, existem duas maneiras de analisar a composição fenomenal (Goff, 2017, p. 296):

- a. Composição restrita, a perspectiva de que apenas alguns conjuntos de objetos são de tal modo organizados que compõem outros objetos, enquanto, por outro lado, há conjuntos de objetos cuja forma os impede de compor um objeto.
- b. Composição irrestrita, a visão de que quaisquer objetos podem ser recombinados em outro objeto.

Para os defensores da Composição Fenomenal Irrestrita, *todas* as partículas que compõem um corpo humano estão ligadas entre si por uma relação fenomenal e, portanto, são capazes de produzir outro sujeito consciente (Goff, 2017, p. 296–297). Quanto àqueles que defendem a Composição Fenomenal Restrita, apenas *alguns* grupos de sujeitos, porém não todos, têm entre si um vínculo fenomenal de tal maneira que produz sujeitos conscientes (Cleve, 1990, p. 216–217). Na visão nageliana,

a composição fenomenal pode ser realizada por qualquer sujeito (Nagel, 2004, p. 43-44). Entretanto, como já apontado, um dado sistema pode não estar organizado de modo a produzir um sujeito experiencial. Diante desse problema, temos que seguir uma entre duas respostas, apresentadas por Goff (2017, p. 298), na linha da Composição fenomenal: a resposta nilista, de que sujeitos nunca se combinam para formar outros sujeitos experienciais; ou a resposta universalista, na qual os sujeitos sempre se combinam para produzir um outro sujeito experiencial. Com base na doutrina do combinacionismo, aqui adotada para examinar a colônia de formigas, caso a perspectiva universalista seja correta, então temos que supor ser a experiência subjetiva partilhável, para daí elucidar de que maneira ocorre a ligação fenomenal.

O Pampsiquismo nageliano afirma que a experiência consciente é algo exclusivo do sujeito, ou seja, pertence a ele e não simultaneamente a outro. A experiência da dor, por exemplo, para um indivíduo, jamais seria sentida e compreendida por alguém da mesma espécie. A dor que um animal ferido experiencia seria incompartilhável. Goff e Roelofs (2020) concordam em parte com esse princípio de exclusividade, no sentido de que certas experiências de um dado sujeito não podem pertencer ao mesmo tempo a outro, como o sentimento de ser “eu”. Neste âmbito, segundo Goff e Roelofs (2020), em sujeitos que estão materialmente conectados, certas experiências seriam fracamente exclusivas. Assim, a partilha fenomenal se torna possível. Mas, para que ela aconteça, o argumento da privacidade fraca precisa ser verdadeiro. De acordo com este argumento, um atributo fenomenal é diretamente conhecido por um sujeito e conhecido apenas indiretamente por qualquer sujeito discreto (Goff & Roelofs, 2020, p. 7; Roelofs, 2019, p. 63).

Desse modo, os autores defendem uma partilha fenomenal fraca, isto é, uma única experiência que pode pertencer a múltiplos sujeitos. A partilha fenomenal ocorreria somente quando estes sujeitos estão sobrepostos - um contendo o outro como parte própria, ou ambos partilhando uma única parte própria como, por exemplo, gêmeos siameses com cérebros conectados por uma ponte neural⁵; ou insetos, como formigas, por meio da comunicação multimodal; ou abelhas, através da dança, informando a localização da comida. Em tais situações, uma determinada experiência seria partilhada gerando uma consciência de grupo.

Tendo em conta o argumento combinacionista, a ligação fenomenal, sugere Goff (2016, p. 181-182), poderia ser idêntica à relação espacial; desse modo, um grupo de componentes físicos, estando espacialmente conectados poderia formar um sujeito consciente. Ora, se imaginarmos diferentes modos de objetos se relacionarem espacialmente, teremos tipos distintos de consciência. Supondo que os insetos sociais são estes objetos, a ligação fenomenal que se dá entre os membros de uma colônia dependerá, do modo como estão conectados espacialmente. Nesse último caso, as formigas encontram-se geralmente dispersas no ambiente, elas não têm uma relação espacialmente estreita como as células somáticas, ou uma ligação anatômica como ocorre com o *D. discoideum* na constituição de um corpo multicelular. A integração entre as formigas é basicamente química. Se elas estivessem estreitamente conectadas no espaço físico,

⁵ Estou me referindo ao célebre caso das irmãs siamesas Tatiana e Krista Hogan que compartilham uma ponte talâmica entre os cérebros. A ponte neural parecia permitir às duas partilharem não somente sinais sensoriais, mas, talvez, pensamentos conscientes. Ver DOMINUS, Susan. *Could Conjoined Twins Share a Mind?* The New York Times Magazine, 2011. Disponível em: <https://www.nytimes.com/2011/05/29/magazine/could-conjoined-twins-share-a-mind.html?pagewanted=all>. Acesso em 02 de fevereiro de 2021.

teríamos uma forte unidade fenomenal; mas se, por outro lado, estão frouxamente conectadas, a unidade fenomenal será fraca. Consequentemente, se levarmos em conta o tipo de arranjo dos constituintes físicos e a unidade fenomenal, então, é natural supor que haja distintas formas de experiência.

Posto isto, Goff, a partir da premissa da Composição Material, propõem que, caso os constituintes básicos estejam dispostos na forma de um organismo temos então uma entidade subjetiva. Entretanto, como ele mesmo admite (2017, p. 298), existe uma dificuldade implícita no argumento da composição, a fronteira entre o orgânico e o não-orgânico, entre o que é verdadeiramente um organismo e o que não é. Tal dificuldade está na imprecisão do arranjo adequado das entidades físicas no momento certo para produzir um organismo. Em um ponto qualquer do tempo, escreve Goff, se pode dizer que há uma situação limite, em que não sabemos determinar o início e o fim de um organismo, um intervalo preciso em que temos um zigoto e antes somente espermatozoide e óvulo. Como resultado disso, se não há um princípio de organização, temos que aceitar a hipótese de casos limítrofes na composição material (Ibid., p. 298).

Neste âmbito, se a colônia de formigas, de acordo com a teoria da organismalidade, é um organismo genuíno, tendo em vista que apresenta alta cooperação entre os membros (Queller & Strassman, 2009, p. 3147-3148), na perspectiva da teoria da autopoiese, por sua vez, não constitui um organismo devido à ausência de uma membrana semipermeável (Thompson, 2013, p. 130). Logo, parece razoável dizer que ela se encontra na fronteira da organismalidade. Aqui, considerando a experiência consciente em entidades vivas, a colônia estaria entre ser um organismo consciente e não-consciente. Mas isso depende, é claro, do tipo de acoplamento entre as partes (Varela, 2000,

p. 67-68). Supondo que a colônia de insetos forme algo próximo de um organismo, do ponto de vista da Composição fenomenal, ela poderia se encontrar em uma zona intermediária. A respeito desse problema, Goff (2017, p. 298) escreve: “Se a existência de um organismo é necessária e suficiente para a existência de um sujeito consciente, e se, às vezes, é vago se temos ou não um organismo, segue que às vezes é vago se temos ou não um sujeito consciente”.

5. Respostas?

Estamos diante, portanto do chamado Problema de Fronteira (*Boundary Problem*) ou Problema de Aninhamento (*Nesting Problem*) (Rosenberg, 2004, p. 81-90; Fekete et al, 2016, p. 1-16; Schwitzgebel, 2014, p. 1-21; Roelofs, 2019, p. 59), que se trata do limite existente entre a experiência das partes individuadas e a experiência do todo individuado. Ou seja, se a mente do sujeito componente desaparece ou não na mente do todo. O Problema de Fronteira induz a afirmar que as experiências do sujeito compartilhadas com cada um de seus componentes podem produzir, ou não, uma combinação genuína. Isso significa dizer que os sujeitos componentes tanto podem ter o mesmo conjunto inteiro de experiências que o sujeito adicional, como também haver uma lacuna entre a experiência das partes e a do todo (Roelofs, 2019, p. 59-60).

Não sabemos, logo, com base nessa perspectiva, se um dado sistema complexo tem uma experiência integrada. A resposta depende mais de como se entende essa organização. Daí o problema de fronteira, a imprecisão em determinar se a experiência dos sujeitos componentes desaparece ou não na experiência de ser o conjunto e também se são contínuas ou descontínuas entre si. Como bem apontou Goff (2017, p. 298), é possível imaginar estruturas biológicas intermediárias, as quais

corresponde, por sua vez, uma consciência indeterminada. Seguindo esse raciocínio, desenvolvemos a resposta à questão título deste trabalho, a partir de uma leitura epistêmica da relação parte-todo. Ou seja, precisa ter uma composição adequada e não sabemos se realmente está de modo a produzir uma unidade fenomenal. Sendo assim, imaginando que ela seja, de fato, um organismo, teremos as seguintes respostas embasadas no problema de fronteira (Rosenberg, 2004, p. 84-85):

R₁.*Cada formiga é um sujeito experiencial, mas a colônia não é.* Esta possibilidade é verdadeira se a partilha fenomenal entre as formigas não contém informação com representação sofisticada ou se a estrutura resultante da cooperação entre os elementos não é capaz de unificação fenomenal.

R₂.*Cada formiga é um sujeito experiencial, assim como a colônia.* Esta possibilidade é verdadeira se o pampsiquismo constitutivo for verdadeiro. Nesse âmbito, a experiência do todo estaria fundada e seria explicada pela experiência de suas partes combinadas em uma forma especial. Precisamente, a colônia estaria organizada de maneira a suportar uma consciência.

R₃.*A colônia seria um sujeito experiencial, mas nenhuma das formigas seria.* Esta possibilidade implica que as mentes componentes seriam anuladas pela mente do todo, ou que, os componentes não são sujeitos experienciais, mas, a partir da relação existente entre eles emerge um sistema capaz de ter uma fenomenalidade.

A posição com a qual o combinacionismo parece estar comprometido é a R₂, uma vez que se alinha ao argumento da composição fenomenal irrestrita. Entretanto, se pensarmos a respeito

da forma como as formigas estão relacionadas no espaço, teremos uma explicação diferente. Por exemplo, em um cérebro dividido (*split-brain*) temos uma consciência desunificada (Nagel, 1979, p. 151-159); em dois cérebros ligados por um único tálamo cortical teremos duas mentes individuadas que, por meio de um canal, compartilham algumas experiências, sem resultar em uma unidade fenomenal (Langland-Hassan, 2013, p. 1737-1756; Roelofs, 2019, p. 11); já na hipótese de uma fusão cerebral, as mentes individuadas no processo de fusão perderiam sua independência gradualmente até serem anuladas e substituídas pela supermente (Fekete et al, 2016, p. 14; Roelofs, 2019, p. 277-288).

Agora, em comparação a ameba *D. discoideum* e ao sifonóforo *P. physalis*, as formigas não estão fisicamente conectadas a ponto de gerar um corpo multicelular. Na verdade, elas estão dispersas no ambiente. Ainda que afirmemos ser o padrão emergente capaz de sustentar uma subjetividade coletiva, a falta de uma conexão estreita pode implicar no caráter experiencial da unidade. Esse detalhe parece relevante, pois põem em dúvida se ela é ou não uma unidade genuína, ou, se está próximo de alcançar a unidade. Neste último caso, a colônia de formigas se encontraria no ponto limite da organicidade. Analisando, pois, os conceitos de ordem espacial das entidades físicas e unidade fenomenal (Goff, 2017, p. 300), imagino uma quarta resposta ao problema:

R₄. *A colônia seria um sujeito experiencial, mas sua experiência é relativa ao arranjo espacial dos componentes.* Esta possibilidade implica que a colônia, mesmo sendo capaz de sustentar uma consciência, seria fracamente integrada, se encontrando em um grau intermediário, ou, em uma zona cinzenta.

Não temos ideia do que é ser uma consciência em uma zona cinzenta porque se trata de um estudo em seus estágios iniciais, porém, se considerarmos que a colônia não tem uma fronteira material (Thompson, 2013, p. 130-131; Varela, 2000, p. 222) e seus elementos estão espalhados, ainda que a forma como interagem seja suficiente para uma unidade fenomenal, é possível que seja fraca. Não apenas por esta razão, mas também porque não temos certeza se o conteúdo informacional partilhado é rico. Em comparação, um organismo multicelular composto por *D. discoideum* poderia ter uma unidade fenomenal maior, devido à proximidade espacial das partes.

Por conseguinte, segundo Roelofs (2019, p. 282), afirmar que a unidade fenomenal é vaga, é dizer que “há um ponto médio na qual não é determinadamente verdadeiro, nem determinadamente falso que a experiência dos componentes seja fenomenalmente unificada, porque suas relações não são diferentemente ricas o suficiente, nem diferentemente não são”. É neste sentido que eu argumento ser impreciso dizer que a colônia de formigas tem uma consciência. Ela se parece com a água-viva, por não ter um cérebro central que conecte as informações sensoriais em uma representação interna. Em outras palavras, é porque funciona como uma rede neural distribuída que a colônia de formigas estaria na fronteira de um sistema genuinamente consciente.

Portanto, se uma colônia de insetos himenópteros sociais está consciente, isso dependerá da característica de sua composição. A resposta, a fim de determinar se um sistema parecido com uma sociedade está apto a ser um sujeito experiencial, requer a solução do problema da combinação, o que permanece, até o momento, insolúvel. Não obstante, o modo como as partes estão conectadas é epistemicamente discutível. Em suma, uma resposta “sim” ou “não” é, por ora, incerta.

Conclusão

Durante o desenvolvimento desse trabalho examinei a hipótese de a colônia de formigas ser um sujeito consciente. Considerei, com base no conceito de composição que, se ela fosse um organismo genuíno, poderia ter algum tipo de consciência. Em vista disso, analisei a colônia de formigas sob o ponto de vista da composição material e fenomenal. Ou seja, argumentei que as formigas, sendo sujeitos experienciais, poderiam reunir suas experiências em um todo unificado. A consciência da colônia, nesse caso, seria contínua com suas partes microconscientes. Contudo, tendo em conta que a unidade experiencial depende do modo como as partes do indivíduo estão relacionadas espacialmente – se próximas ou espalhadas; haveria, pois, um limite na composição de uma consciência. Isso significa que organismos genuínos, ou próximos de se tornarem organismos, apresentariam diferentes tipos de consciência fenomênica.

No caso da colônia de formigas, a maneira como as formigas estão distribuídas no espaço, implicaria no nível de unificação fenomenal. É evidente que há objeções, como afirmar que apenas a integração da informação, por meio de sinais químicos, por exemplo, seria suficiente, não importa se os membros estão estreitamente próximos ou espalhados. Porém, o modo como as partes estão dispostas na constituição do corpo reflete no nível de consciência do sujeito. Por conseguinte, conclui que a colônia de formigas teria uma fraca integração das experiências individuais. A consciência da colônia se encontraria em uma região intermediária. Ora, tendo em conta que a colônia talvez se encontre em uma linha divisória da organicidade, ela também estaria em uma zona cinzenta da fenomenalidade.

Posto isto, suponho que a incerteza quanto à experiência fortemente unificada em uma colônia de formigas parece coerente. As outras três respostas baseadas na relação parte-todo também são válidas. Tanto é possível uma continuidade quanto uma descontinuidade entre a experiência das partes e a do todo. Entretanto, acho a quarta resposta, que afirma ser a colônia de formigas um caso de consciência limítrofe, mais apropriada. Essa resposta leva em conta os dois lados da composição: o físico e o fenomênico. Ambos são conceitos imprecisos na teoria pampsiquista. Em função disso, a colônia ter uma consciência imprecisa não parece absurdo. Penso, com base nisso, que a consciência pode ser uma propriedade indeterminada da organização.

Enfim, a R_1 , R_2 , R_3 e R_4 podem servir de paradigma para examinar outras formas de colônias organizadas, considerando a interação, comunicação e a disposição das partes. Não é possível garantir que todas se encontrem em um contexto análogo ao da colônia de formigas, mas, em comparação, afirmar que algumas tenham um grau maior de consciência, enquanto outras, ausente.

Referências

- BLOCK, Ned. Troubles with Functionalism. In: CHALMERS, David J. (Ed.). *Philosophy of Mind: classical and contemporary readings*. New York, Oxford: Oxford University Press, 2002, p.94-98.
- CHALMERS, David J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996, 432 p.
- CHALMERS, David J. Consciousness and Its Place in Nature. In: CHALMERS, David J. (ed.). *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press USA, 2002, p. 247-272.
- CHALMERS, David J. *Panpsychism and Panprotopsychism*, 2013, p. 1-32. Disponível em: <http://consc.net/papers/panpsychism.pdf> . Acesso em 20 de fevereiro de 2021.

CHALMERS, David J. *The Combination Problem for Panpsychism*, 2017, p. 1-37. Disponível em: <http://consc.net/papers/combination.pdf> . Acesso em 20 de fevereiro de 2021.

CLEVE, James Van. *Mind – Dust or Magic?* Panpsychism versus Emergence. *Philosophical Perspectives*, Vol. 4, 1990, p. 215-217.

FEKETE, Tomer; VAN LEEUWEN, Cees; EDELMAN, Shimon. *System, subsystem, hive: Boundary problems in computational theories of consciousness*. *Frontiers in Psychology*, 2016, Vol. 7, Art. 1041, p. 1-17. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01041/full>. Acesso em 05 de outubro de 2019.

GOFF, Philip. *Consciousness and fundamental reality*. New York: Oxford University Press, 2017, 290 p.

GOFF, Philip. *The Phenomenal Bonding Solution} to the Combination Problem*. In L. Jaskolla (ed.), *Panpsychism: Contemporary Perspectives*. New York: Oxford University Press. 2017, p. 283-302.

GOFF, Philip; SEAGER, William; ALLEN-HERMANSON, Sean. Panpsychism. In: ZALTA, Edward N. (ed.). *The Stanford Encyclopedia of Philosophy*. California: Stanford University, 2017. Disponível em: <https://plato.stanford.edu/archives/sum2020/entries/panpsychism>. Acesso em 25 de outubro de 2020.

GOFF, Philip. *Galileo's Error: Foundations for a New Science of Consciousness*. New York: Pantheon Books, 2019, 256 p.

GOFF, Philip; ROELOFS, Luke. In Defence of Phenomenal Sharing. In: BUGNON, Julien; NIDA-RUMELIN, Martine (Eds.). *The Phenomenology and Self-Awareness of Conscious Subjects*. Routledge, 2020. Disponível em: https://www.philipgoffphilosophy.com/uploads/1/4/4/4/14443634/in_defence_of_phenomenal_sharing_v1.pdf. Acesso em 2 de maio de 2021.

KLEIN, Colin; BARRON, Andrew B. Insects have the capacity for subjective experience. *Animal Sentience*, Vol. 1, Issue 9, N° 1, 2016, p. 1-54. Disponível em: <https://www.wellbeingintlstudiesrepository.org/animalsent/vol1/iss9/1/>. Acesso em 03 de agosto de 2020

- LANGLAND-HASSAN, Peter. Introspective Misidentification. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. vol. 172, no. 7, Springer, 2015, p. 1737–58. Disponível em: <http://www.jstor.org/stable/24704179>. Acesso em 14 de novembro de 2021.
- MALLATT, Jon; FEINBERG, Todd E. Insect consciousness: Fine-tuning the hypothesis. *Animal Sentience*, Vol. 1, Issue 9, Nº 10, 2016, p. 1-16. Disponível em: <https://www.wellbeingintlstudiesrepository.org/animsent/vol1/iss9/10/>. Acesso em 03 de agosto de 2020.
- MERKER, Björn. Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behavioral and Brain Sciences*. Vol. 30(1), 2007, p. 63-81. DOI: 10.1017/S0140525X07000891.
- NAGEL, Thomas. *Mortal Questions*. Cambridge: Cambridge University Press, 1979, 226 p.
- NAGEL, Thomas. *Visão a Partir de Lugar Nenhum*. Trad. Silvana Pereira. 1ª ed. São Paulo: Martins Fontes, 2004, 405 p.
- QUELLER, David C; STRASSMANN, Joan E. Beyond society: the evolution of organismality. *Philosophical Transactions of the Royal Society Biological Sciences*. Vol. 364, Issue 1533, 2009, p. 3143-55.
- ROELOFS, Luke. *Combining Minds: A Defence of the Possibility of Experiential Combination*. Dissertation, University of Toronto, 2015, 420 p.
- ROELOFS, Luke. *Combining Minds: How to think about composite subjectivity*. New York: Oxford University Press, 2019, 336 p.
- ROSENBERG, Gregg. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. New York: Oxford University Press, 2004, 325 p.
- SCHWITZGEBEL, Eric. If materialism is true, the United States is probably conscious. *Philosophical Studies*, Vol. 172, p.1697–1721, 2015. DOI: <https://doi.org/10.1007/s11098-014-0387-8>.
- SCHWITZGEBEL, Eric. *Are garden snails conscious? Yes, no, or *gong**. *The Splintered Mind*, 2018. Disponível em: <http://schwitzsplinters.blogspot.com/2018/09/are-garden-snails-conscious-yes-no-or.html>. Acesso em 3 de agosto de 2020.

SØVIK, Eirik; PERRY, Clint. The evolutionary history of consciousness. *Animal Sentience*. Vol. 1, Issue 9, Nº 19, 2016, p. 1-5. Disponível em: <https://www.wellbeingintlstudiesrepository.org/animsent/vol1/iss9/19/> Acesso em 03 de agosto de 2020.

STRASSMAN, Joan E.; QUELLER, David C. The Social Organism: Congresses, Parties, and Committees. *The Society for the Study of Evolution*. Vol. 64, Nº 2, 2010, p. 605-616.

STRASSMAN, Joan E.; QUELLER, David C. How social evolution theory impacts our understanding of development in the social amoeba *Dictyostelium*. *Development, Growth & Differentiation*, Vol. 53, 2011, p. 597-607. Disponível em: <https://doi.org/10.1111/j.1440-169X.2011.01272.x> Acesso em 18 de agosto de 2020.

THOMPSON, Evan. *A Mente na Vida: Biologia, Fenomenologia e Ciências da Mente*. Lisboa: Instituto Piaget, 2013, 542 p.

VARELA, Francisco. *El fenómeno de la vida*. Santiago de Chile: Dolmen, 2000, 472 p.

WHEELER, William M. *The ant-colony as an organism*. *Journal of Morphology*. Vol. 22, 1911, p. 307-325.

2

WHAT IS IT LIKE TO BE AN ANT COLONY?

Luke Roelofs

If a colony of ants had a mind, what sort of mind might it have? In particular, what sort of *conscious experience* might a colony have - not just the individual ants, but the colony itself, as a collective whole? More broadly, what can we say about the minds - and particularly the conscious experiences - of sophisticated collectives with simple members? In this paper I discuss this question with particular focus on ants, because ants are the iconic eusocial animal, and the most overwhelmingly successful and widespread lineage thereof. But what I say likely applies also to other eusocial insect colonies, like those of termites and some species of bees and wasps. And it could be at least a starting point for thinking about forms of artificial swarm intelligence that could arise in the future. What is it like to be a hive mind?

Of course, it's possible the answer is: nothing, because hives don't really have minds. You might think this follows from a general principle about collectives - that their members can be conscious, but the collective itself can't. Or you might think this is a contingent truth about present-day insect colonies: whatever the necessary conditions are for consciousness, ant nests, bee hives, and termite swarms don't meet them. You could combine this with a similar thought about individual ants and other insects: that they're just little automata, mindlessly going about their work without any of it feeling like anything *for them*. Or you could, alternatively, think that it's precisely the consciousness of individual ants and bees that poses a barrier to

their colonies being conscious: something about being made up of conscious parts is incompatible with being conscious as a whole.

These objections to hive minds turn on much broader questions about what consciousness is, what it takes to be conscious, and how consciousness fits into the natural world. Elsewhere I've argued against all of these sorts of scepticism (Roelofs Ms.), and will summarise those arguments only briefly. I note that there are many existing proposals of functional criteria for consciousness that are met by individual insects like ants (see, e.g., Klein; Barron, 2016; Barron; Klein, 2016; Feinberg; Mallatt, 2013, 2016; Bronfmann *et al.*, 2016, Ginsberg; Jablonka, 2019; Birch *et al.*, 2020; Malatt *et al.*, 2021). I argue moreover that if we apply these functional criteria in an open-minded way, there is a good chance that some eusocial insect colonies meet them also (compare arguments in Schwitzgebel, 2015, 2020; Friedman, Sovik, 2021; Fonseca, 2022), partly in virtue of the functional capacities of their individual members and partly in virtue of the way their members interact. I argue moreover that we have no good reason to adopt any form of 'Anti-Nesting Principle', to the effect that conscious parts cannot make up a conscious whole (cf. Putnam, 1965; Oizumi, *et al.*, 2014; Schwitzgebel, 2015): such principles have implausible results and serve no useful theoretical purpose. On the contrary: I have in other work advocated 'combinationism' (Roelofs, 2019), the idea that conscious minds can be constituted by other conscious minds interacting in suitable ways. Combinationism is obviously congenial to the possibility of hive minds, and my hope is that the ensuing discussion of eusocial insects may help illustrate the fruitfulness of combinationism as an outlook.

But my primary question here is not *whether* ant colonies could have conscious minds, but the conditional: if they did, how should we think about those minds? Obviously, we cannot hope for much

confidence or completeness here: famously, humans struggle to really grasp what it's like even to be a different human, and face major obstacles imagining what it's like to be other animals. So how much less can we expect to imagine what it's like to be a group of many thousands of tiny insects? But we can still hope for empirically and theoretically informed guesses about the broad outlines and major structures. For instance, in a famous discussion of the elusiveness of consciousness, Thomas Nagel writes:

We cannot form more than a *schematic* conception of what it is like [to be a bat]. For example, we may ascribe *general types* of experience on the basis of the animal's structure and behavior. Thus we describe bat sonar as a form of three-dimensional forward perception; we believe that bats feel some versions of pain, fear, hunger, and lust, and that they have other, more familiar types of perception besides sonar. But we believe that these experiences also have in each case a specific subjective character, which it is beyond our ability to conceive (Nagel 1974, p. 439, emphasis added).

Here Nagel distinguishes between the unanswerable questions - what specifically does bat sonar feel like? - and the more tractable project of sketching out the rough contours of a creature's experience: a 'schematic conception' of some 'general types' of experience and their relations. A bat may perceive the world through different senses, but we can still reasonably believe that it has perceptual experiences, caused by stimuli impacting parts of its body, and that these are distinct from but coupled with its other experiences - 'pain, fear, hunger, and lust', for example, though no doubt human hunger is only an imperfect guide to the specific subjective character of bat hunger.

My goal in asking about ant colony minds is just to get a sense of what their rough contours might be, and how they are likely to differ from the rough contours of bat minds, human minds, and even individual ant minds. In a spirit of empirically-guided speculation, the

next six sections will consider the following six questions, and suggest the following preliminary answers:

1. Does an ant colony have a brain or nervous system? I'll argue that it has a distributed nervous system with no central brain, consisting of the nervous systems of all its members and the media through which they transmit information to one another.
2. Does an ant colony have a body? I'll argue that its body consists of both the bodies of all its members, together with their nest and eggs.
3. How does the colony's mind relate to the minds of its component ants? I'll argue that it has a distinct mind that is nevertheless nothing over and above theirs, in a way that is best understood through the idea of *shared* token experiences with *divergent* structure.
4. What does an ant colony perceive? I'll argue that ant colonies likely perceive, consciously or unconsciously, those stimuli which their members both perceive and communicate about, and that the form of their perceptions is likely dictated by how much information about stimuli their members can communicate.
5. What other kinds of experience does an ant colony have? I'll argue that it likely does not have cognitive or imaginative experiences, but rather experiences affect, volition, and perception in a systematically connected way, without the relatively clean differentiations that structure human experience.

6. What is the moral significance of ant colony consciousness? I'll argue that although the colony may have a degree of moral status in virtue of its consciousness, this moral status is best viewed not as some extra value added to the value of its members, but as simply their value considered collectively.

1. An Open-Air Brain

One natural worry about the idea of collective consciousness is that collectives don't seem to have their own brain, or more broadly their own nervous system (cf. List 2018, pp.305-308). But is this really true? Obviously, there's no extra brain sitting deep inside the nest, but there is plenty of neural tissue in the colony, biologically healthy and working to coordinate the system's behaviour in light of information available to it. Isn't that what a brain is, and what a brain does?

Let's try to sharpen up this idea. If we were looking for something that functions as an 'organ of coordination', we might as a first pass say that the ant colony's brain is simply the totality of its members' brains. After all, the colony's behaviour is coordinated and intelligent partly in virtue of the member's behaviour being so, so let's look to what coordinates the behaviour of each individual ant: its brain, aka cerebral ganglia). But this by itself is only part of the story: the ants' collective behaviour depends crucially on the interactions between them, in particular on the ways that they transmit information. Just as a brain would not be a brain without the synapses that connect neurons with each other, so the colony would not be a colony if the ants could not communicate (see esp. Gordon, 2010). So, we should think of the colony's brain as the totality of ant brains together with the physical connections that transmit signals between them. Some of these

connections involve other parts of the ants, like the motor nerves by which one ant drums their antennae on another, and the sensory nerves by which the other detects the drumming. But some connections involve stuff outside any ant's body, like airborne pheromones, or the pheromone trails they lay down to guide each other.

This 'brain' is quite different from a human brain, or indeed an ant's brain, in many ways. For a start, it is not biologically insulated: it is not all in one place, surrounded by protective tissue. It is, so to speak, an 'open-air brain'. For another thing, it is not centralised: rather than being all in one place, it is distributed throughout the colony's members. In effect, the colony reaches out to hunt for food, fight enemies, and gather resources, using its brain. If we did things this way, virtually any sort of bodily damage would involve brain damage. Yet the colony's brain differs from ours in other ways that mitigate the vulnerability of grabbing things with grey matter. Whereas our individual neurones stand fixed in specific relations to each other, the colony's neurones come in little packages (ant brains) which move constantly relative to one another¹. This likely prevents them from developing the sort of exquisitely fine-tuned organisation that our brains develop through our years of learning, but with the corresponding advantage that it makes them much more robust against the destruction of individual ants. Our brains famously have a degree of 'neuroplasticity', such that in response to damage they can re-organise themselves and press surviving structures into new use to replace those that have been lost. But we can do this only slowly and imperfectly,

¹ Of course, if they move too far and become unable to return to the colony, these neurons will cease to be a functional part of the colony's brain, just as the ant containing them ceases to be a functional part of the colony, and a piece of brain tissue removed from my head ceases to be a functional part of my brain.

whereas a colony which routinely loses many of its members must be able to do it as a matter of course². In sum, the colony could be said to have a de-centralised, hyper-plastic, open-air brain.

Based on these differences, of course, we might question if this really deserves to be called a brain. If we apply the term ‘brain’ narrowly, we might say that it’s not really a brain, just as an ant, or a hypothetical android, doesn’t literally have a ‘brain’ (instead they have cerebral ganglia, or a central processor, respectively). But we shouldn’t get caught up on terminology here: even if an ant or android lacks a brain strictly speaking, they have something that plays the same role: a physiological structure of coordination that provides a plausible physical supervenience base for whatever mental life they might have.

2. The Body Politic

A related question about collective consciousness in general is whether groups have bodies. Can you consciously feel, sense, or move if you have no body? Would hive minds have to be in some sense disembodied? At least for ant colonies I would suggest the opposite: their minds are hyper-embodied. Because while it’s true that an ant colony doesn’t have a body separate from the bodies of its members, that doesn’t mean it doesn’t have a body.

Here we have to step back and ask: what makes a particular part of the physical world ‘my body’? It seems to be a cluster of factors:

² The colony is not perfectly plastic: the experiences of older ants seem to be somewhat important, so it seems likely that losing enough of them might degrade the colony’s collective intelligence more than losing an equal number of younger ants. See discussions of colony-level memory in, e.g. Gordon 2018a, 2018b; cf. Franks *et al.*, 2007.

- My direct sensitivity to things happening to it (not based on detecting something else that indicates them);
- My ability to move its organs directly (not by first moving something else connected to it);
- The supervenience of my experiences on states of it³.

So, to identify the body that belongs to the colony's mind, we might look for whatever can be directly felt by, can be directly moved by, and contains the physical basis of, its mind. Is there such a candidate body? It's not immediately clear: certainly, the body of an individual ant doesn't seem like the right candidate. So, we must ask a more basic question: what makes a certain part of the physical world 'a body', of the sort that could be mine? What makes all the different parts of my body into parts of a single body? Again, it seems natural to point to a clustering of many factors, particularly:

- Being tightly coupled, in that each is physically attached to another part, and hard to remove (unlike a garment or a handheld tool);
- Having a common origin: each was produced by the growth, secretion, division, etc. of another part; and (as a product of this fact) sharing a genetic code.

These factors need not be present in every case to the same degree. Some cells in my body have no genetic code at all (like my red blood cells), and some people are genetic chimeras with multiple genetic

³ For discussion of the forms and grades of embodiment - including cases where these factors come apart, see esp. Dainton, 2008, p. 201-235, cf. Dennett, 1978.

codes; a foetus growing in my womb is coupled to me and grown from me but is genetically distinct, and so might be seen either as a body part or as a separate body contained within mine.⁴ And a prosthetic limb might be more or less strongly coupled to the rest of my body - potentially being as hard to remove as any other limb - while having a fully external origin, making its status as a body part contested⁵.

So, let's take a body to be a system of (at least somewhat) causally coupled parts, with (at least somewhat) connected origins, that can be controlled by a single mind. I would then say: the totality of individual members of an ant colony, together with their eggs and their physical nest, can be viewed as a single body in this sense (cf. Friedman *et al.* 2020). The various parts of this body are genetically linked due to their common origin: they are all offspring of the queen, or products produced by the queen's offspring. The nest might be considered somewhat like a mollusc's shell, or like the casing that a caddisfly larva forms around itself from pebbles and other detritus; plausibly these are borderline cases of body parts⁶. Indeed, for leafcutter ants who farm fungus deep within their nest, these fungi might be considered as genetically-different parts of the same body, rather like the bacteria making up human gut flora (cf. Gilbert *et al.*, 2012).

Of course, the notable thing about this 'body' is that the coupling of its parts together is much looser than the coupling of my different tissues. The ants are physically separated, and move independently, so if they form a body, it is a 'distributed body'. But there is nevertheless an important form of coupling here: the ants interact with each other,

⁴ For discussions of the metaphysics of pregnancy, see Kingma 2018, Finn, 2021.

⁵ For discussion of the status of prostheses, see Aas, 2019.

⁶ Different sorts of nests might qualify more or less strongly as products of the insects who live in them: nests made of wax, paper, etc. qualify more than those which are simply dug into the earth.

and with their nest, in ways that help ensure that each tends to congregate with the others, return to the others, follow the others, and thus form a whole that ‘hangs together’ consistently across time.

If the distributed mass of ants, eggs, and nest, can be viewed collectively as ‘a body’, then it seems well-suited to be ‘the body’ of whatever colony mind might be present. For it contains within it what I above suggested we could view as the colony’s nervous system: the assemblage of individual ants’ nervous systems, together with the signals they transmit to each other via pheromones, touch, etc. Moreover, it seems natural to think that any action the colony takes must involve movement of some part of this body. Likewise, whatever the colony senses, it senses only by some individual ants being affected. Thus, this distributed body seems to fulfil the conditions to embody the colony.

One interesting result is that ant colony minds are likely to be *more* embodied than ours are. Since our brains are present in only one part of our body, we can internally entertain all sorts of thoughts, form all sorts of images, even hallucinate all sorts of perceptions and actions, regardless of what is going on with the rest of our body. Counting on our fingers can help us do maths, but we can also do maths ‘in our heads’ - binding our hands does not stop us counting.⁷ By contrast, an ant colony’s brain is spread throughout its body, and the ways that information flows among ants will be heavily influenced by their spatial distribution. As a result, even if our cognition is to some extent embodied, an ant colony’s cognition is likely ‘hyper-embodied’ relative

⁷ For discussion of the extent to which human cognition is embodied, see e.g. Gallagher, 2005; Shapiro, 2007; Clark, 2008.

to ours, in that its particular form and content depends a lot on the shape and disposition of this body.

3. Shared Experiences with Divergent Structure

The last two sections argued that while an ant colony does have a brain and body, these largely consist of the interacting brains and bodies of its members. This gives some reason to expect that its mind will also, in some sense, just consist of the interacting minds of its members. Attributing a mind to an ant colony doesn't mean adding a separate mind 'on top of' the minds of the ants: it's more like zooming out and seeing that their minds, considered together, constitute a single mind. This mind will be 'new', in the sense that it is genuinely a mind, and isn't identical to any of the ant's individual minds. (To put it another way, the colony is a conscious subject, distinct from any individual ant, and so is a 'new' subject of consciousness.) But the colony mind is still, as philosophers like to say, 'nothing over and above' the suitably related ant minds, just as a brain is nothing over and above the connected neurones that compose it. Every property of the whole derives from the properties of the parts, and the way they are put together: nothing appears out of nowhere⁸.

But how should we think about this constitutive relationship? In *Combining Minds*, I suggested the following first-pass principle:

Conditional experience inheritance: If a part of X has an experience which plays a certain functional role, then X shares that experience, with the same or similar functional role, to the extent that the part in question is connected

⁸ Strictly speaking, one might point out that the colony's brain contains some parts that aren't part of any ant's brain (the signals and media connecting the ants). But insofar as the informational complexity of those signals taken in isolation, is vastly less than that of an ant brain, taken in isolation, it still seems approximately right to say that the mind of the colony is nothing over and above the coordinated minds of its ants.

to the rest of X such that its experience has sensitivity to, control over, and coordination with other events occurring in X (Roelofs 2019, p. 183).

This principle is schematic, in that it leaves open how to understand sensitivity, control, and coordination in any given case. It is easiest to understand in cases like a human being and their brain, where one specific part of a whole seems like a candidate for being ascribed all the same experiences. My experiences (and thoughts, judgements, and other mental states) supervene on events in, and thus might plausibly be thought to belong to, my brain. But since they also belong to me, and I am not my brain, this has been thought to generate a puzzle of “too many thinkers”⁹. Combinationism, however, suggests there is nothing wrong with both me and my brain sharing the very same experiences, because as well as being a part of me, there is a systematic sensitivity of events in my brain to events in the rest of me (due to my sensory nerves and other input channels) and a systematic control by events in my brain of events in the rest of me (due to my motor nerves and other output channels). Because the experiences of my brain satisfy these functional conditions with respect to me as a whole, I inherit them from my brain. We don’t need to agonise about whether it is really the whole organism of the brain that is conscious: once we recognise this functional relationship, we can say “both” and move on.

Things get more complicated when there are many different parts of a whole which it could inherit experiences from. No single ant is like my brain: none of them is the ‘leader’, monitoring and directing what the rest do (the queen, in particular, is better seen as the colony’s reproductive system, not its nervous system - a gonad, not a brain). Rather, each ant’s experiences satisfy the functional conditions laid down

⁹ For discussion see, e.g. Merricks, 2001; Blackmon, 2016, 2021; Roelofs, 2022.

by Conditional Experience Inheritance to a small degree: they are sensitive to other events going on in the colony, though only somewhat, and they exercise control over other events in the colony, though only somewhat. This is where the third functional condition - coordination - becomes important. One ant's experiences are only very imperfectly sensitive

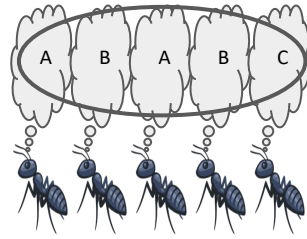


Figure 1. Each ant's experience (the thought bubbles) comprises a set of experiential elements (represented by letters). The colony's experience (the large oval) incorporates some of these elements, giving it a very different total experiential field from any individual ant.

to broader changes in the colony-body, but as we consider larger subsets of ants, their sensitivity increases, both because they are statistically more likely to pick up on things happening to the colony, and because they effect a valuable form of feedback: one ant detecting a change is likely to signal that change to others, who can signal to others, potentially amplifying the signal so that the whole colony is alerted. Similarly, one ant's experiences have only a very small influence on what the colony body as a whole does, but they can exert greater control if they cause that ant to act in ways that alert or recruit enough other ants to join in a given activity.

The implication of these considerations, I would suggest, is that ant experiences are inherited by the colony not individually but *en masse* (as schematically indicated in Fig. 1). In the colony's mind, what registers as 'a single experience' is the blended experiences of many individual ants, so that the colony has no phenomenological sense of what is coming from one ant or from another, any more than we have a phenomenological sense of how much of a given experience is due to one neurone or to another. Individual ant's experiences are

“phenomenally blended” together for the colony (cf. Roelofs, 2014), so that although the colony’s conscious field is entirely constituted by experiential elements shared with individual ants, there is a radical difference in the overall pattern of relations that organises those shared elements. After all, the colony is both a very different beast than any of its members, and yet at the same time it largely just is them.

For example, suppose that many ants all register a dangerous shortage of stored food (perhaps in virtue of shifts in the pattern of which other ants they interact with at what frequencies, cf. Gordon, 2010), and release pheromones that have the effect of shifting more ants towards foraging roles. From each ant’s perspective, the experience of detecting the shortage occurs together with many others, like monitoring their own individual location, seeing other ants ahead of them, reading local pheromone trails, etc. - and the shortage-detection-experiences of the other ants do not appear in this perspective. But from the colony’s perspective, these individual details fade out, since they do not connect to any larger patterns, while the many lack-of-food-detection-experiences, though they occur in different ants, cohere as a single experience because they reinforce each other’s common causes and effects - they cohere phenomenally because they cohere functionally. But this does not require adding a new experience for the colony - its experience of hunger can be assembled from those of individual ants. The same experiential raw materials take on different overall patterns at different scales, because of the different informational and functional structures operative at those scales¹⁰.

¹⁰ A clarification: talking of the ‘elements’ of a total conscious state always faces a tricky ambiguity, between two readings of ‘element’ (the same is sometimes true about talk of ‘experiences’ as countable things). There is a constitutive reading, on which an experience has two elements if there are two experiences which could have occurred separately but which together add up to it, and there is also a phenomenological reading, on which an experience has two elements if it presents its subject with two

4. Seeing Through 100,000 Eyes

Let's illustrate how the idea of conditional inheritance works by looking at one particular sort of experience: perception. If ant colonies are conscious, what external things are they conscious of - what do they sense? The principle of Conditional Experience Inheritance, discussed above, suggests one way of approaching this question: start with the perceptual experiences of individual ants (if any), and then ask which of those (if any) fulfil the conditions of sensitivity, control, and integration, relative to the whole colony. Fortunately, work on what individual ants (and bees, wasps, and termites) can perceive is rich and abundant: we are learning a lot about eusocial insect perception, in particular of light and odour¹¹. The tricky part is evaluating the inheritance conditions: how much of this realises the functional role of perceptual experience on the collective level? We can divide this into three more specific questions¹²:

- Of the information collected by ants, how much, if any, is available to influence larger-scale colony-level dynamics?

distinguishable things, two contrasting objects or qualities. These senses need not correspond: an experience might have two elements in the constitutive sense which blend seamlessly rather than appearing contrastively, or might have two elements in the phenomenological sense while being metaphysically simple. Indeed, it might have two elements in both senses, without the two lining up: constitutive elements A and B might blend seamlessly into an experience as of a contrast between X and Y. In this section I am suggesting that constitutive elements are shared between ant and colony, potentially yielding very different arrays of phenomenological elements.

¹¹ See, for instance, Bhagavan and Smith 1997, Alem *et al.* 2016, Franks *et al.* 2007, Collett 2010, Galizia *et al.* 2012, and especially Chittka, 2022.

¹² There is a (limited) analogy here with parallel questions about nerve cells and the human body: how much information from human distal nerves reaches the brain? What brain processing constitutes perception? And which perceptions are conscious rather than unconscious? But of course, there is one glaring disanalogy, namely that humans perceive through distal nerves which must relay information to the central nervous system. Insofar as the colony is just made up of ants, the ants are simultaneously acting like sensory nerve cells (e.g. in the eye or skin) and like nerve cells in the brain.

- Of the information made available at the colony level, what perceptual states - mental representations of external conditions - does it constitute, if any?¹³
- Of the colony-level states that represent external conditions, which, if any, are phenomenally conscious?

These questions are much more difficult to investigate than the question of what individual ants can detect; they are beset with both conceptual and empirical challenges. The first, about what information is transmitted between ants so as to flow through the colony, requires not only identifying the mechanisms by which ants communicate (e.g. pheromone trails, antenna-drumming), and identifying what information these mechanisms actually transmit (e.g. simple ‘friend-vs-foe’ identification, a more or less intense ‘follow me’, or spatial coordinates¹⁴), but then also cataloguing the overall patterns of which ant interact with which, what information they transmit, how much of that gets passed on, and to whom. The second and third questions then require correct theories of perceptual representation and perceptual consciousness, respectively, and moreover require a sense of how to determinately apply those theories to a very differently-constructed organism. These questions are extremely hard, for both practical and conceptual reasons, so sheer ignorance looms large here.

¹³ An interesting related question is: does the colony have sense organs? Arguably it does in a sense: it has all the eyes, antennae, etc. of its members. But in another sense, it seems not to: all parts of its body serve both sensory and motor functions.

¹⁴ My impression of the current consensus is that bees do exchange relatively rich information through their ‘waggle dance’, whereas ants do not: they lay down pheromone trails that say something like ‘follow me’, and when meeting they can identify each other as colony-mates, as injured or infected, etc. but don’t describe food sources beyond this. See Reznikova and Ryanko, 1994, 2011 for disagreement.

However, we do have at least one well-studied example of eusocial insect communication where we have a fairly strong grasp of what information is represented and how that information flows: the dances of honeybees, especially when used to select a new nest location (Lindauer, 1985; Seeley, 1992, 1995, 2003; Seeley; Burman, 2001). Bees returning from foraging or exploration perform a dance whose direction, duration, and intensity encode the direction, distance, and value of a food source or potential nest site; other bees who observe the dance can follow these ‘instructions’ to find the site and, upon returning, perform their own dance. In the process of nest-selection, numerous scouts find and report various candidate sites, inducing other bees to investigate those sites; eventually, the most attractive sites accrue more and more visitors, who dance longer and more energetically, until the best site has all or almost-all of the bees dancing for it; when this consensus is achieved, the colony departs *en masse* to build a nest there.

Following Huebner (2014, p. 192-199), I take this to be a compelling example of a perceptual representation at the collective level. The dances coding for a particular site are, collectively, action-guiding; they have a clear evolved function; they have a clear external-world referent; and they are unambiguously collective, because the colony can compare multiple locations even if each bee only visits one. But are they phenomenally conscious? Here I see no particular hope for consensus, so let us simply observe that if any perceptual process of a eusocial insect colony is conscious, the nest-selection process seems like a promising candidate, due to its recurrent, interactive, and global character.

So far, I’ve characterised the colony’s perceptions in terms of their referent - the external properties that they track. But of course, this is only half the story; if we are interested in the possibility of colony

consciousness, we will wonder how these external properties are represented in the colony's perception. Just as the physical structure of wavelengths underdetermines the qualities of colour vision, so the physical facts about a nest site underdetermine the qualitative character of the honeybee colony's perceptual experience of it. And while the exact quality of that perceptual experience is presumably unimaginable to us, we might reasonably think that the dimensions of its quality space will be shaped by the way that information is transmitted between bees. For instance, an individual scout bee might perceive a nest site in detail, getting a sense of its temperature, height above the ground, colour, size, shape, etc. But to the extent that all of this is summarised into a single dimension of 'nest site quality' that determines the duration of its dance, the colony's perception of that site might be expected to be phenomenologically much simpler, characterised only by a degree of quality, a distance, and a direction¹⁵.

Suppose that we postulate consciousness in the nest-selection process: how do other processes look by comparison? If we associate consciousness with being sufficiently *global*, as on 'global workspace' theories (see, e.g., Baars, 2017; Dehaene *et al.*, 2017), then most other processes will seem less promising than nest-selection - even the everyday use of the waggle-dance to allocate foragers to nectar and pollen sources. Here the same representational code (direction, duration, intensity, etc.) is used, but since each nectar site becomes depleted by a few visitors there is not the 'winner-takes-all' feedback structure of nest-selection (it would be bad for every bee to end up at the same flower). Insofar as each dance transmits information from one

¹⁵ Things might get more complicated than this - the perception of the best site might be phenomenally different when all returning scouts have the same quality estimate than when they vary, for instance, if that makes a functional difference. Everything depends on the structure of information processing.

bee to some small number of others, it might be closer to unconscious processing, or more like the less-attended visual periphery than focal vision. Perhaps the honeybee colony has a mind that spends most of the year in what we would view as ‘unconsciousness’, or just a sort of inattentive flow of many habitual actions, but rouses itself to focal consciousness once a year to make the decision with the highest stakes.

One significant difference between bees and ants is that bees fly to and from their food, while ants walk. Walking allows for the laying down and later reading of pheromone trails, and it seems that much of the communication among ants depends on such trails. This sort of communication seems much less centralised, both spatially and temporally, and to that extent seems an even weaker candidate for consciousness by global workspace standards. But much depends on how we individuate representations: on the smallest scale, the information at any particular point on a particular trail is decoded ‘one step at a time’, by one ant at a time, but if we consider the distribution of pheromones as a whole, as a single ‘map’, then we might say that it is globally accessible in that it can be ‘read’, albeit slowly, by the colony, using any or all of its ‘consuming systems’ (foraging ants). So, evaluating how “globally” anything is processed is hard to do with any confidence: this underlines that the global workspace theory (and its rivals), becomes less and less determinate as we remove more and more of the architectural features that characterise the human brain (as argued by Carruthers, 2019; for discussion see Birch, 2020, Michel, 2021). But we might at least suggest that if there is phenomenal consciousness here, it is likely to be more diffuse than that associated with the honeybees’ dances, less a matter of crisp structured images and more a matter of slow unguided feelings-out. Perhaps bee swarms are perceptually conscious and ant colonies are not; or perhaps bee swarms

have experiences closer to vision or hearing, while ant colonies have experiences closer to smell or touch.

We should also remember that perception is not exclusively of the outside world; organisms also perceive their own internal state, and in the case of ant colonies this means in particular perceiving the distribution of members among tasks and locations within the nest. Although some progress has been made in identifying the different sorts of chemical signals used to mark out areas and coordinate task allocation (see e.g. Heyman *et al.*, 2017), it is harder to decode the informational structure of such processes or identify their representational content, if any. For all we know, there might be stronger candidates for conscious colony-level perception here than in external foraging behaviour: ant colonies might be more aware of their own bodies than of their surroundings. But it is worth stressing that the colony's perception, particular as embodied in its pheromone trail network, seem to be capable of presenting many objects and their relations within a spatial environment (like individual insects and vertebrates do), rather than merely responding to the presence or absence of immediate stimuli (like jellyfish and plants do). In that sense, they might be thought to have conscious sensory 'images' - though perhaps more like our body image than like a visual image - particularly in how they interweave sensory with motor dimensions, as the next section discusses.

5. Feeling Antsy

Although perceptual consciousness is arguably the most well-studied form of consciousness, it is not the only form. We humans undergo, at a minimum, emotional/affective and imaginative

experiences, as well as being conscious of desires, decisions, thoughts, intentions, memories, and others (though there is room for disagreement about which of these states are or are not conscious in the phenomenal sense). Which of these might an ant colony have? If we take the presence of functional structure at the colony level as our guide, we should probably consider it unlikely for colonies to have thoughts, imaginings, future intentions, or the active recall of specific memories - that is, mental states that internally represent absent objects. These would require processes that are apt to guide and direct behaviour at a colony-wide level, but do not immediately issue in action - representations of specific objects or situations that are held in (collective) mind and not outwardly expressed. It seems unlikely that we will find structures with this combination of sophistication and permanence among the constantly circulating ants.

A more promising thing to look for is consciousness connected with action and motivation - feelings, emotions, desires, and the experience of intentional movement itself. Let's start by asking what actions the colony takes - in a sense of 'action' broad enough to allow for animals generally, including insects, to take various actions like walking or biting¹⁶. The paradigmatic form of action among most animals is locomotion, movement of the whole body from one place to another, and there are some instances where ant colonies do this. Army ant colonies, most notably, travel across the forest floor every day, and the above-mentioned example of honeybees selecting and then flying to a new nest site would also qualify. However, for the most part this is

¹⁶ I take it as a plausible premise that, as Frankfurt puts it, "The conditions for attributing the guidance of bodily movements to a whole creature, rather than only to some local mechanism within a creature, evidently obtain outside of human life", since they are met "when a spider moves its legs in making its way along the ground" (1978, p. 162).

not how eusocial insect colonies operate: to put it figuratively, they tend to remain in one location while extending their members, like tentacles, out to explore the environment and gather food, like a housebound octopus. In this sense, we should not require whole-body locomotion as a necessary condition of action. But of course, we do not want to count anything that any ant does as an action of the colony, any more than anything one of my cells does is my action. As a first pass, we might attribute a given ant-behaviour to the colony in proportion to both 1) how many of the colony's ants are involved, and 2) how far their behaviour is best explained collectively, by the functioning of communication systems that work to coordinate ant behaviour towards colony goals. By this, admittedly rough, standard we would say that when a single ant is separated from the colony, what it does is not meaningfully attributable to the colony, whereas what a single ant does *because of* a mechanism allocates ants to tasks (e.g. cleans the larvae, sets off to explore) is somewhat attributable to the colony. When many ants successively navigate to and harvest from a food source by following each other's pheromone trails, that is substantially attributable to the colony, and when all of the ants drop what they were doing and shift to evacuation or defence duties because alarm pheromones are flooding the nest in response to invasion, that is very strongly attributable to the colony.

So perhaps we can make sense of 'what the colony does', but how far does any of this make it into consciousness? Is there, for instance, an experience of desire, leading to an experience of volition, leading to an experience of performing a certain action? I think the most likely answer is 'no' - not because there is no experience at all, but because the steps are likely not differentiated in this way. For us, desire exists earlier than, and need not lead to volition, which exists earlier than, and need not lead

to, action, and even the experience of action is potentially distinct from the actual execution of action. But this is because distinguishable brain areas are dedicated to these different functions. Insofar as the colony's brain is made up of ant brains and the signals they exchange, such a differentiation seems unlikely: there are no boss ants with the job of deciding what other ants should do. So, if there is any consciousness associated with what the colony does, our best human model might be cases where desire, willing, and action all occur together in a seamless succession - perhaps when the impulse to stretch an uncomfortable limb arises in unbroken tandem with the process of stretching.

This point can be pushed further: the colony likely does not have a sharp differentiation between states that purely direct action and states that purely describe reality. It might rely entirely on what Millikan (1995) calls 'pushmi-pullyu representations'. After all, the colony's distributed nervous system mingles motor and sensory nerves, since each ant has both: hence its perceptions might generally also be actions, and vice versa. In our brains we have whole areas devoted to largely 'receptive' processing: taking in information to detect how the world is. And we also have whole areas devoted to largely 'active' processing: identifying goals, planning actions, and coordinating the physical implementation of an action plan. Hence, we are used to having some mental states (perceptions, beliefs, etc.) that are focused on telling us how things actually are, and others (desires, intentions, etc.) that are focused on how we want to change them. But the ant colony only has ants, and each ant has both sensory and motor systems. So, we might conjecture that its

mental states will often have a fused sensorimotor character, simultaneously detecting the world and directing responses at it¹⁷.

This applies to the type of action that I above noted as perhaps the most fully attributable to the colony itself: the rapid shift to defensive behaviour in response to attack. As Millikan notes, alarm signals are plausible examples of pushmi-pullyu representations: they at once convey descriptive information (“danger! We’re under attack!”) and prompt actions (fleeing, fighting, evacuating eggs or young, blocking nest entrances, etc.). But they also show a striking resemblance to emotional responses in individual organisms like us: a systematic shift in how resources are distributed, what sorts of stimuli are salient, and what actions are prepared for. Perhaps it is useful to think of colonies as experiencing things like the ‘fear’ of being attacked and scrambling defences, the ‘contentment’ of being full of food and shifting more workers away from foraging towards nest maintenance, the ‘hunger’ of finding little food and extending more scouts in more new directions, as long as we remember that these feelings are not distinct from the perceptions that prompt them and the actions that they prompt.

This leads to the following suggestion: ant colonies might experience affective states in a more vivid and centralised way than they experience anything else, including sensory images (if they experience those at all). Their consciousness may thus be heavily dominated by feelings over images, by comparison with the consciousness of animals like us (and ants) that is more dominated by large, dedicated, sense organs.

¹⁷ Millikan (1995, p. 190) cites the waggle dance of the honeybee as a prime example of a pushmi-pullyu representation, though of course having this character relative to the bees who produce and receive the representation does not automatically entail that it has this character relative to the colony.

6. Non-Additive Moral Status

Consciousness is often thought to be a sufficient condition for moral status: if a being can experience, then it seems natural to think it has interests that we should give at least some weight to in our moral deliberations. Does ant colony consciousness imply something like this for the ant colony? Does it matter, morally?

On the assumption that consciousness grants moral status, ant colony consciousness would indeed imply ant colony moral status. But there is an important caveat; its value may not be something to add to whatever value its component ants have (for recent discussions of the moral status of arthropods, see esp. Birch *et al.*, 2020; Sebo, 2023). Just as its consciousness is nothing over and above theirs, the moral consideration it demands may turn out to be just another way of viewing the moral consideration that they demand. That doesn't mean that the colony doesn't matter: it means that if it matters and its members matter, these might be two sides of the same coin.

Whether this is true depends on both how exactly we understand the relation between consciousness and value link, and how exactly we understand the relation between ant-consciousness and colony-consciousness. In particular, it's common to think that the most morally important sorts of consciousness are valenced states, which present something as *good* or *bad*. Examples include suffering, joy, distress, contentment, and perhaps desire and frustration. Certainly, these are the forms of consciousness that most directly relate to the content of a conscious being's interests. Above, I argued that the consciousnesses of the colony and its ants exhibit shared elements with divergent structure. This raises the question whether the sort of valence that matters for moral status has more to do with the elements

of consciousness (that are shared) or with the structure (which diverges). When my experience is pleasant, is that because it contains a specific pleasant element (e.g. a pleasant taste in my mouth), or because of how that pleasure affects my attention, efforts, or the overall organisation of my consciousness?¹⁸

On one way of thinking, valence is atomistic: the valence of a complex experience is a linear sum of the amount of positive valence in the ingredients that went into it (e.g. particular pleasant sensations), minus all the negative (e.g. particular unpleasant sensations). On this way of thinking, it might make sense to associate valence with the shared elements, and therefore to attribute completely non-additive moral status to the colony, since whatever valence is present in its experiences can only be recycled from what is present in the ants' experiences.

Alternatively, we might think that valence is to a greater or lesser degree holistic, depending on how one's total conscious field is organised. In *Combining Minds* (p. 4 ff), I argue that holistic properties like this are division-relative, and can differ between a whole and its parts even if the whole is nothing over and above the parts. This opens up the possibility that the valence in the colony's consciousness could diverge from that in the ants' - e.g. that it might feel good in virtue of them feeling bad. There would then be genuinely different moral interests, that would have to be balanced against each other.

However, even if valence is holistic and such divergence is possible in principle, there is little reason to think that it is actual. There are few circumstances where a colony benefits from things that distress most of its ants - there may be many cases where a single ant's painful self-

¹⁸ For a fuller discussion of this issue as it arises across various forms of experiential sharing, see Roelofs and Sebo, Ms.

sacrifice serves the colony, but only because it serves the interests of the many other ants. The colony just doesn't seem organised enough to systematically pursue interests that conflict with the drives of its members. So, it appears likely that even if the colony is conscious, and even if consciousness is sufficient for moral status, we do not need to add the colony to our moral calculations as a new entry: any calculation that factors in the welfare of insects, including ants, will have thereby already included that of their societies¹⁹.

Conclusion

In an interview in 2005, Daniel Dennett said “What is it like to be an ant colony? Nothing, I submit, and I think most would agree intuitively”. Perhaps so, but in other work I have defended the idea of ant colony consciousness, and in this discussion, I've taken it as a starting conjecture. Specifically, I maintain that the ant colony might well have a conscious mind which, although it is distinct from that of any individual ant, and has its own characteristic structure and phenomenology, reflecting the functional structure of information processing at the collective level, nevertheless is nothing over and above the interacting minds of the ants, just as the colony is nothing over and above the interacting ants.

My aim here hasn't been to suggest final answers, but to organise and connect questions. The question of what an ant colony consciously

¹⁹ Might something like this apply also to human brains and either the social groups they form or the neurones that compose them, if either of the latter were to be or become conscious? This is an interesting question, considered more fully in Roelofs and Sebo Ms., but one obvious fact is that the difference in cognitive sophistication and integration between a human brain and either a society or a neuron is enormous, especially compared to the difference between an ant and an ant colony (where it is, in fact, hard to say confidently which is more sophisticated). Consequently, on a holistic view of valence it is much more likely that the human brain's mental life would involve valuable states with no counterpart in either the minds of either its parts or its groups.

perceives depends partly on what ants can perceive, partly on what information ants communicate to each other and in what patterns, and partly on what the right theory is of the structural requirements for conscious perception. And the question of what other sorts of experiences an ant colony might have will likewise depend on difficult questions both myrmecological (how does information flow within the colony, and how does that coordinate ant behaviour?) and philosophical (what is it to feel emotions, or desires, or to consciously act?). As a first pass, I've suggested that ant colonies probably sense their surroundings, and the quality of food sources laid out in them, but do so in a slow and decentralised way, while they probably experience assaults on or disturbances in their next much more immediately and rapidly, though still on a slower timescale than is typical for individual conscious animals - simply because it takes more time for information to travel between ants than between neurones. I've also suggested that an ant colony's awareness of its world is more like our body image than like visual images: extended, and complex, but dynamic and embodied, with a systematic interweaving of sensation, affect, and action. But all of these are just attempts at educated guesses, and my aim is just to convey what sort of guesses, educated by what sort of scientific investigations, we should be attempting. It is extremely hard to know what it's like being either an ant colony or an ant, but we can work towards at least a schematic conception.

References

AAS, Sean. Prosthetic embodiment. *Synthese*, v. 198, n. 7, p. 6509-6532, 2019.

ALEM, S., PERRY, C. J., ZHU, X. F., LOUKOLA, O. J., INGRAHAM, T., SOVIK, E., CHITTKA, L. Associative mechanisms allow for social learning and cultural transmission of string pulling in an insect. *PLOS Biology*, v. 14, n. 10, 2016. Epub: e1002564.

- BAARS, Bernard. The global workspace theory of consciousness: Predictions and results. In: S. SCHNEIDER, S.; VELMANS, M. (Eds.). *The Blackwell Companion to Consciousness*. Hoboken, NJ: Wiley-Blackwell, 2. Ed., p. 227–242, 2017.
- BARRON, Andrew; KLEIN, Colin. What insects can tell us about the origins of consciousness. *PNAS*, v. 113, n. 18, p. 1-9, 2016.
- BHAGAVAN, S.; SMITH, B. H. Olfactory conditioning in the honey bee, *Apis mellifera*: effects of odor intensity. *Physiology and Behavior*, v. 61, p.107–117, 1997.
- BIRCH, Jonathan. Global Workspace Theory and Animal Consciousness. *Philosophical Topics*, v. 48, n. 1, p. 21-38, 2020.
- BIRCH, Jonathan. The search for invertebrate consciousness. *NOÛS*, v. 56, p. 133– 153, 2022.
- BIRCH, Jonathan; GINSBURG, Simona; JABLONKA, Eva. Unlimited Associative Learning and the origins of consciousness: a primer and some predictions. *Biology & Philosophy* v. 35, n. 56, p. 1-23, 2020.
- BIRCH, Jonathan; SCHNELL, Alexandra; CLAYTON, Nicola. Dimensions of Animal Consciousness. *Trends in Cognitive Sciences*, v. 24, n. 10, p. 789-801, 2020.
- BIRCH, Jonathan; BURN, Charlotte; SCHNELL, Alexandra; BROWNING, Heather; CRUMP, Andrew. *Review of the Evidence of Sentience in Cephalopod Molluscs and Decapod Crustaceans*. London: Department for Environment, Food and Rural Affairs, 2021.
- BLACKMON, James. Hemispherectomies and Independently Conscious Brain Regions. *Journal of Cognition and Neuroethics*, v. 3/4, p. 1-26, 2016.
- BLACKMON, James. Integrated Information Theory, Intrinsicity, and Overlapping Conscious Systems. *Journal of Consciousness Studies*, v. 28, n. 11–12, p. 31–53, 2021.
- BRONFMAN, Zohar; GINSBURG, Simona; JABLONKA, Eva. The Transition to Minimal Consciousness through the Evolution of Associative Learning. *Frontiers in Psychology*, v. 7, p. 19-54, 2016.

CARRUTHERS, Peter. *Human and Animal Minds: The Consciousness Questions Laid to Rest*. Oxford University Press, 2019.

CHITTKA, Lars. *The Mind of a Bee*. Princeton University Press, 2022.

CLARK, Andy. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford, New York: Oxford University Press, 2008.

COLLETT, Matthew. How desert ants use a visual landmark for guidance along a habitual route. *Biological Sciences*, v. 107, n. 25, p. 11638-11643, 2010.

DAINTON, Barry. *The Phenomenal Self*. Oxford University Press, 2008.

DAWSON, E. H.; AVARGUES-WEBER, A.; CHITTKA, L.; LEADBEATER, E. Learning by observation emerges from simple associations in an insect model. *Current Biology*, v. 23, n. 8, p. 727–30, 2013.

DEHAENE, Stanislav; LAU, Hakwan; KOUIDER, Sid. What is consciousness, and could machines have it? *Science*, v. 358, p. 486-492, 2017.

DENNETT, Daniel. *Where Am I? Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press, 1978.

DENNETT, Daniel. *What do you believe is true even though you cannot prove it?* Entrevista com The Edge: Annual Question, 2005. Disponível em: <https://www.edge.org/response-detail/11902>. Acesso em: 12 maio de 2023.

FEINBERG, T. E.; MALLATT, J. The evolutionary and genetic origins of consciousness in the Cambrian Period over 500 million years ago. *Frontiers in Psychology*, v. 4, n. 667, 2013.

FEINBERG, Todd E.; JON, M. Mallatt. *The Ancient Origins of Consciousness: How the Brain Created Experience*, MIT Press, 2016.

FINN, Suki. Methodology for the metaphysics of pregnancy. *European Journal for Philosophy of Science*, v. 11, n. 3, p. 1-19, 2021.

- FONSECA, Anderson. *A Colônia de Formigas está Consciente? Uma Análise sob a ótica do Pampsiquismo*. Dissertação (Mestrado em Filosofia). Programa de Pós Graduação de Filosofia e Métodos, Universidade Federal de São João del-Rei - MG, 2022.
- FRANKFURT, Harry. The Problem of Action. *American Philosophical Quarterly*, v. 15, n. 2, p. 157-162, 1978.
- FRANKS, Nigel; HOOPER, James; DORNHAUS, Anna; AUKETT, Philippa; HAYWARD, Alexander; BERGHOFF, Stefanie. Reconnaissance and latent learning in ants. *Proceedings of the Biological Sciences*, v. 22, n. 274 (1617), p. 1505-1509, 2007.
- FRIEDMAN, D. A.; JOHNSON, B. R.; LINKSVAYER, T. A. Distributed physiology and the molecular basis of social life in eusocial insects. *Horm Behav*, v. 122, n. 10, p. 47-57, 2020. DOI: 10.1016/j.yhbeh.2020.104757. Epub PMID: 32305342, 2020.
- FRIEDMAN, Daniel A.; EIRIK, Søvik. The Ant Colony as a Test for Scientific Theories of Consciousness. *Synthese*, v. 198, p. 1457-1480, 2021.
- GALIZIA, C. G., EISENHARDT, D.; GIURFA, M. (eds.). *Honeybee Neurobiology and Behavior: A Tribute to Randolph Menzel*. Dordrecht: Springer, 2012.
- GALLAGHER, Shaun. *How the Body Shapes the Mind*. Oxford, Oxford University Press, 2005.
- GILBERT S. F.; SAPP, J; TAUBER, A. I. A symbiotic view of life: we have never been individuals. *Q. Rev. Biol.*, v. 87, n. 4, p. 325-341, 2012. DOI: 10.1086/668166. PMID: 23397797.
- GINSBURG, Simona; JABLONKA, Eva. *The evolution of the sensitive soul: learning and the origins of consciousness*. MIT Press, Cambridge, 2019.
- GODFREY-SMITH, Peter. Gradualism and the evolution of experience. *Philosophical Topics*, v. 48, n. 1, p. 201-220, 2020.
- GORDON, Deborah. *Ant encounters: interaction networks and colony behavior*. Princeton University Press, 2010.
- GORDON, Deborah. Ant Colonies Retain Memories That Outlast the Lifespans of Individuals. *The Smithsonian Magazine*, 12. dez. 2018. Disponível em:

<https://www.smithsonianmag.com/science-nature/ant-colonies-retain-memories-outlast-lifespans-individuals-180971022/>. Acesso em: 12 nov. 2022.

GORDON, Deborah. The Ecology of Collective Behavior in Ants. *Annual Review of Entomology*, v. 64, n. 3, p. 1–16, 2018b.

HEYMAN, Yael; SHENTAL, Noam; BRANDIS, Alexander; HEFETZ, Abraham; FEINERMAN, Ofer. Ants regulate colony spatial organization using multiple chemical road-signs. *Nature Communications*, 2017. DOI: 10.1038/ncomms15414.

HOFSTADTER, Douglas. Prelude...Ant Fugue. In: GODEL, Escher. *Bach: An Eternal Golden Braid*. Basic Books, p. 149–201, 1979.

HUEBNER, Bryce. *Macro cognition: A Theory of Distributed Minds and Collective Intentionality*. Oxford University Press, 2014.

KEY, Brian; ZALUCKI, Oressia; BROWN, Deborah. Neural Design Principles for Subjective Experience: Implications for Insects. *Frontiers in Behavioral Neuroscience*, v. 15, 658037, 2021.

KLEIN, Colin; BARRON, Andrew. Insects have the capacity for subjective experience. *Animal Sentience*, v. 9, n. 1, 2016.

KINGMA, Elselijn. Lady Parts: The Metaphysics of Pregnancy. *Royal Institute of Philosophy Supplement*, v. 82, p. 165–187, 2018.

LINDAUER, M. The dance language of honeybees: the history of a discovery. *Experimental Behavioral Ecology*, Eds. B. Hölldobler, M. Lindauer, p. 129–40, 1985.

LIST, Christian. What Is It Like to Be a Group Agent? *NOÛS*, V. 52, n. 2, p. 295–319, 2016.

MALLATT, Jon; BLATT, Michael; DRAGUHN, Andreas; ROBINSON, David; TAIZ, Lincoln. Debunking a myth: plant consciousness. *Protoplasmia*, v. 258, p. 459–476, 2021.

MENZEL, R., GREGGERS, U., SMITH, A., BERGER, S., BRANDT, R., BRUNKE, S., BUNDRÖCK, G., HÜLSE, S. *et al.* Honey bees navigate according to a map-like spatial memory. *Proceedings of the National Academy of Sciences*, v. 102, p. 3040–3045, 2005.

- MERKER, Bjorn. The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition*, v. 14, p. 89-114, 2005.
- MERKER, Bjorn. Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, v. 30, p. 63-81, 2007.
- MERRICKS, Trenton. *Objects and Persons*. Oxford: Oxford University Press, 2001.
- MICHEL, Matthias. Review of Peter Carruthers, *Human and Animal Minds: The Consciousness Questions Laid to Rest*. *Philosophical Review*, v. 130, n. 4, p. 619-623, 2021.
- MILLIKAN, Ruth. Pushmi-pullyu representations. *Philosophical Perspectives*, v. 9, p. 185-200, 1995.
- OIZUMI, M.; ALBANTAKIS, L.; TONONI, G. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLOS Computational Biology*, v. 10, n. 5, 2014.
- PUTNAM, H. Psychological predicates. In: W. H. Capitan & D. D. Merrill (Eds.), *Art, mind, and religion*. Liverpool: University of Pittsburgh Press, 1965.
- REZNIKOVA, Z., and RYABKO, B. Experimental study of the ants' communication system with the application of the Information Theory approach. *Memorabilia Zoologica*, v. 48, p. 219-236, 1994.
- REZNIKOVA, Z.; RYABKO, B. Numerical competence in animals, with an insight from ants. *Behaviour*, v. 148, p. 405-434, 2011.
- RELOFS, Luke. Phenomenal Blending and the Palette Problem. *Thought: A Journal of Philosophy*, v. 3, p. 59-70, 2014.
- ROELOFS, Luke. *Combining Minds: How to Think about Composite Subjectivity*. Oxford University Press, 2019.
- ROELOFS, Luke. No Such Thing as Too Many Minds. *Australasian Journal of Philosophy*, v. 102, n. 1, p. 131-146, 2024. DOI: <https://doi.org/10.1080/00048402.2022.2084758>.

- ROELOFS, Luke. (Ms). *Anty-Nesting: Why Eusocial Insect Colonies are Plausibly Conscious*.
- ROELOFS, Luke; SECO, Jeff. (Ms.) *Overlapping Minds and the Hedonic Calculus*. *Philosophical Studies*, v. 181, p. 1487-1506, 2024.
- SCHWITZGEBEL, Eric. *If Materialism Is True, the United States Is Probably Conscious*. *Philosophical Studies*, v. 172, p. 1697-1721, 2015.
- SCHWITZGEBEL, Eric. *Is there something it's like to be a garden snail?" Philosophical Topics*, v. 48, n. 1, p. 39-64, 2020.
- SEBO, Jeff. *The Rebugnant Conclusion: Utilitarianism, Insects, Microbes, and AI Systems*. *Ethics, Policy & Environment*, 2023. DOI: 10.1080/21550085.2023.2200724.
- SEELEY, T. *The tremble dance of the honey bee: message and meanings*. *Behavioral Ecology and Sociobiology*, v. 31, p. 375-383, 1992.
- SEELEY, T. *The Wisdom of the Hive*. Cambridge: Harvard University Press, 1995.
- SEELEY, T. *Consensus building during nest-site selection in honey bee swarms: the expiration of dissent*. *Behavioral Ecology and Sociobiology*, v. 53, p. 417-424, 2003.
- SEELEY, T., and Buhrman, S. *Nest-site selection in honey bees: how well do swarms implement the 'best-of-N' decision rule?* *Behavioral Ecology and Sociobiology*, v. 49, p. 416-427, 2001.
- SHAPIRO, Lawrence. *The Embodied Cognition Research Programme*. *Philosophy Compass*, v. 2, n. 2, p. 338-46, 2007. DOI: 10.1111/j.1747-9991.2007.00064.x.
- STROEYMEYT, Natalie; GRASSE, Anna; CRESPI, Alessandro; MERSCH, Danielle *et al*. *Social network plasticity decreases disease transmission in a eusocial insect*. *Science*, v. 362, n. 6417, p. 941-945, 2018.
- TYE, Michael. *Are insects sentient?* *Animal Sentience*, v. 1, n. 9, 2016b.
- VARELA, Francisco, Thompson, Evan, and Rosch, Eleanor. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press, 1991.

3

TOWARDS A SCIENCE OF CONSCIOUSNESS AND SOCIAL COMPLEXITY ... FOR ANTS

Maria Luiza Iennaco

Daniel Friedman

1. Systems, Measurements, and Theories in Consciousness Science

Understanding and measuring consciousness is a perennial challenge for scientists, theorists, and practitioners (such as doctors and social workers) alike. This quest to develop a better understanding of the correlates and causative factors of consciousness is motivated by various significant implications for multiple fields (Ledoux; Michel; Lau, 2020; Seth, 2021).

In this article, we set some context for consciousness studies, in terms of scientific pluralism (Longino, 2000), specifically regarding the target System, empirical Measurements, and utilized Theories. After reviewing several leading formal theories of consciousness, we highlight recent work that uses Active Inference to consider the material, cognitive, and experiential aspects of conscious awareness in diverse systems. Then, with an Active Inference lens in place, we consider how recent multidimensional frameworks on animal consciousness (Birch; Schnell; Clayton, 2020) can be operationalized to apply to the eusocial Hymenoptera (all ants, some bees, and wasps) as global systems of interest.

In pursuit of expanding this human-centered approach to consciousness studies, without falling into the trap of a “view from nowhere” (Nagel, 1989), broadly we see the challenge is to find an approach to measure and model conscious systems in a bottom-up

fashion, describing and predicting heterogeneity among systems (in practice) using a common principle to enable comparability. With that in mind, here we focus on how emerging physics-based approaches such as Active Inference and the Free Energy Principle may help us understand and model such heterogeneity. To provide an empirical grounding for this work and point towards specific examples of available measures and experiments, we highlight the ant colony system.

1.1 Systems

Ants are a globally distributed clade of insects that obligately live in multi-nestmate colony system (Kass *et al.*, 2022). Beyond their fascinating and economically important biology, ants have also stimulated inquiry into diverse areas such as distributed systems design (Mullen; Monekosso; Barman; Remagnino, 2009; Heylighen, 2016) and philosophy (Hofstadter, 1999; Gordon, 1992; Costa, 2002; King, 2006). The thousands of ant species around the world vary widely in almost every conceivable nestmate and colony-level trait (Linksvayer, 2015), summarized by a variety of measures (Holland; Bloch, 2020; Traniello; Linksvayer; Coto, 2022).

The collective behavior of ant colonies, represented by their niche modifications and interaction patterns (Gordon, 2010), has long been compared to the human brain (Hofstadter, 1999, Chapter 11, *Prelude... Ant Fugue*). Analogous to various measures of cognitive complexity in human brains and societies, insect researchers have proposed “social complexity” measures in ants, encompassing “many individual and colony-level traits and characteristics such as colony size, polymorphism, and foraging strategy” (Anderson; McShea, 2001). As we will develop here, we believe that ant colonies are appropriate systems for considering a plurality of

bottom-up and multi-measurement methods for studying consciousness, amidst the plural theoretical avenues, as scaffolded and scoped by Active Inference and the Free Energy Principle.

To extend consciousness studies beyond humans, here focusing on the ant colony system, we approach this inquiry from the perspective of Measurements & Theories – which is to say in terms of the data collected in the empirical process (measurements) as it relates with the cognitive frame of the investigator as it guides their analysis (theories). Specifically through applied Active Inference, we sketch a mapping and relationship between the phenotype of social complexity in ants and the phenotype of consciousness across systems. We will suggest that insights from the study of social complexity in ants may shine some pragmatic as well as epistemic light on current debates in consciousness studies.

1.2 Measurements

The concept of consciousness is highly ambiguous and multifaceted (Barúss, 1987; Tassi; Muzet, 2001; Velmans, 2009; Rosenthal, 2009; Friedman; Søvik, 2019). Philosophers and researchers have developed a range of measurement and modeling techniques for consciousness based on the idea that consciousness itself (among other cognitive features) has measurable and quantifiable types and degrees, both regarding its states (arousal) and its content (awareness). Some of the literature on consciousness in humans, for instance, operationalizes the idea that conscious awareness has linear, hierarchical levels or degrees (e.g. Glasgow coma scale, the Richmond Agitation-Sedation scale, Coma Recovery Scale-Revised [Tindall, 1990; Bayne; Hohwy; Owen, 2016]). However, it is worth noting that these

medical heuristics are designed for clinical utility in limited settings, not as generalized epistemic instruments.

Even though these forms of measuring consciousness are useful in their respective niches, there are conceptual problems in application, generalization, and transfer among the various expressions of consciousness. This non-integration of perspectives within and across approaches exacerbates, and fails to leverage, gaps among the disciplines that propose to study it.

Subjective aspects of consciousness are commonly measured by phenomenology, for example relying on verbal reports from subjects describing their experiences in experiments. These reports are analyzed in conjunction with other data, including neural correlates of consciousness (NCCs), which are brain regions or bodily processes potentially correlated with specific conscious experiences. In this regard, there are some neuroscientific tools used to measure brain activity, neural correlation, and connectivity associated with different levels and aspects of consciousness, such as neuroimaging and eye tracking.

In philosophy, phenomenal consciousness is described as the sensation of being an organism – an intrinsic, ineffable, and private sensation to each one. Such a description places the phenomenon in a fundamentally intellectualist category, dependent on interior experience and limited to human beings since it would be exceedingly difficult to study something with such characteristics in another species or system (Dennett, 1992; Ji, 2023). Meanwhile in medicine, the leading criteria used by scientific and clinical literature, perhaps for reasons such as pragmatic utility and simplicity, usually consider the healthy neurotypical adult human as the maximally or prototypically conscious creature, which can fail to recognize, or even diminish, characteristics of neurodivergent consciousnesses, especially if such

measures are applied outside of the clinic. These gaps or mismatches open room to misguided presumptions and theories that are different in their assessments and precisions, in estimates about the young, unhealthy, non-human, and non-biological manifestations of consciousness. Amidst this setting, broader explorations of consciousness can proxy as or for discussions like “is this system a culpable or praiseworthy adult human?” (Skinner, 1971; Sapolsky, 2023).

After conditioning a line of inquiry upon the system of interest (e.g. a human for a medical clinician, ant for a myrmecologist, bee for an apidologist, etc.), the space of measurements may be fairly constrained by the empirical setting. For example, the human medical researcher might only have access to biosensory data, and the myrmecologist might only have access to the measures made and experiments possible in the field. With that in mind, the following sections briefly review several formal frameworks for consciousness that might even be able to approach extended non-human systems such as ants, in terms of accommodating empirical measurements while grounded in physical theoretical first principles. As models are both constrained and embodied by their constructs, like maps that only capture some aspects of the territory they represent (Cordes; Friedman; Maron, 2020; Andrew, 2021; Ramstead; Sakthivadivel; Friston, 2022), it is to be expected and welcomed that various models of consciousness differ in their considered measurements as well as explanatory and predictive scope.

1.3 Theories

Here we explore more carefully several contemporary empirical theories (here used in a similar sense to frameworks or models) for consciousness. We highlight these theories as examples of how

researchers are currently approaching the study of consciousness in human and non-human systems today. Models of consciousness serve as powerful tools for studying consciousness. The utility of having different models of consciousness is their ability to address distinct functions and processes. When systematically compared, they can operate as allies rather than competing models (Graziano *et al.*, 2020; Del Pin, 2021; Storm *et al.*, 2024).

Nevertheless, we argue that these models are often limited by an idealized conceptualization of consciousness and multiple biases and misguided intuitions. To overcome these limitations, a resolution of multiple biases may be achieved through investigating and defining the minimum criteria of consciousness, as Crick and Koch did in the 1990s (Crick; Koch, 1990), or by identifying common properties that organisms possess as criteria for consciousness, as Seth *et al.* developed in 2005 (Seth; Baars; Edelman, 2005). However, we contend that these efforts will not meaningfully advance consciousness studies if the conceptualization of consciousness continues to rely on intuitive top-down approaches.

Cognitive neuroscientific models of consciousness such as the Global Workspace Theory (GWT) (Baars, 1988; Baars, 2017) simulate neural networks or connections that correspond to or constitute essential aspects of consciousness. GWT posits that consciousness emerges from the activity of multiple brain networks that integrate and disseminate information across the system. These networks function as modules that process information unconsciously in parallel and interact with each other through a global workspace that serves as a gateway to conscious experience. Therefore, consciousness is the outcome of global access to selected pieces of information in the system. This model and its subsequent versions (Dehaene; Changeux, 2011;

Bressler; Kelso, 2016) are widely accepted in the field of consciousness studies and have some empirical support. However, a notable limitation of this approach is the reliance on the specific neuroanatomical structures of the human brain for the existence of consciousness. Although the model was devised and implemented for humans, this reliance may pose a challenge to the description of consciousness in non-human architectures of consciousness.

Informational models of consciousness, such as the Integrated Information Theory (IIT), attempt to formally specify the informational, statistical, mathematical, or otherwise analytical variables that constitute the phenomenal experience (Tononi, 2004; Tsuchiya; Taguchi; Saigo, 2016; Negro, 2022). Based on axioms (Crick; Koch, 1990), the IIT measures the complexity of the phenomenal experience through the variable Φ (phi), which indicates the amount of integrated information within a system. In this model, informational properties are key intrinsic aspects of a system, which will causally influence its state in the past, present, and future, affecting the system qualitatively. Therefore, IIT's postulates aim to determine whether a system is conscious and in what manner. Although many authors who appreciate and emphasize mathematical rigor in theories of consciousness find this approach appealing, the IIT faces some difficulties. To name a few, what does it exactly imply to equate the phenomenal experience to the level of integrated information in the system? Furthermore, given that this theory does not restrict Φ to human beings, is the computation of the value of integrated information, or any quantitative descriptor for that matter, in other systems enough to infer what kind of conscious experience they are having?

This raises the question of whether it is necessary to replace all attempts to measure consciousness in degrees, or if it is more effective to

alter philosophical assumptions and underlying theories, which could meaningfully alter the way any given quantitative measure is interpreted. There are comprehensive examples of studies that have abandoned top-down approaches and embraced a system-agnostic approach to studying consciousness. Starting with bottom-up quantum models of consciousness, some argue that consciousness emerges from the entanglement and superposition of subcellular neural quantum structures (Hameroff; Penrose, 2014; Hameroff, 2021). Proponents of this approach claim that properties of consciousness cannot be accounted for by the paradigm of Newtonian mechanics alone. In some of these models, consciousness is distinct from the “mind”, which would have an independent role from the neural circuits in shaping conscious experience. While these quantum models of consciousness have faced criticism from scientists and philosophers alike (Tegmark, 1999; Derakhshani *et al.*, 2022; Koch; Hepp, 2006; Bools, 1990; Grush; Churchland, 1995; Putnam, 1995) for relying on contentious philosophical assumptions (Atmanspacher, 2020), discussion and development on these quantum approaches continue to this day. Alternative “top-down” quantum cognitive formulations rest not in the material properties of microtubules, but rather in terms of the nature of interfaces and discrete cognitive states. For example, Ramsted *et al.* (2023) structured an “inner screen” model of consciousness that follows the Free Energy Principle in both classical and quantum formulations (Ramstead *et al.*, 2023). Specifically for this “inner screen” model in the ant context, it is interesting to consider physio-chemical interactions among nestmates as holographic screens, passing information throughout the colony analogously to communication among neurons (Hofstadter, 1999).

Besides the quantum world, a variety of other frameworks for modeling conscious states have been proposed, such as that of Tsuchiya

et al. (2016) employed Category Theory to assess the relationship between consciousness and Integrated Information Theory (IIT) (Tsuchiya; Taguchi; Saigo, 2016). Similarly, Safron (2020) combined IIT and Global Neuronal Workspace Theory (GNWT) in a bottom-up fashion based on the Free Energy Principle and Active Inference (FEP-AI) framework (Safron, 2020). Rorot (2021) investigated previous attempts to integrate models of consciousness into Bayesian approaches, which also operate in a bottom-up fashion, to develop what he called the Minimal Unifying Model of consciousness (Rorot, 2021).

Our intention here has been to review without pre-judging the epistemic or pragmatic utility of any of these theories for the ant colony system and beyond. Next, we start building our proposal using a bottom-up approach, towards a broadened view of consciousness.

2. Consciousness in Light of Organism-Environment Co-Determination

Exploring the evolutionary and developmental origins of consciousness is an increasingly popular path among researchers seeking to understand the plurality of its expression, including e.g. co-homeostasis (Ciaunica *et al.*, 2021) and basal cognition (Bongard; Levin, 2023). In this regard, a good starting point may be to discuss what makes conscious cognitive agents co-determinant while distinct from their environment, as this maintained partition may provide a starting point for what we call consciousness, in relationship or as arising from the broader field of mortal computing¹ (Ororbia; Friston, 2023).

In order for an organism to maintain its structural and functional integrity as a locally thermodynamically delimited entity, it needs to

¹ Mortal computing means that “the mathematical calculations/processes that underlie information processing in a biological or artificial system are inseparable from the physical substrate that implements and executes them” (Ororbia; Friston, 2023, p. 3).

fight against global (informational, physical) equilibria, thus remaining distinct from its environment (Schrödinger, 1992; Stewart; Gapenne; Di Paolo, 2010). This, in turn, requires the limitation of the entropy of a thing's own constituent states, and thus temporarily ensures its local thermodynamic equilibrium – this can be viewed as a system “fighting” against the Second Law of Thermodynamics (Ramstead; Badcock; Friston, 2018). In this sense, living beings are defined as thermodynamically open and far-from-equilibrium autopoietic systems. That is, systems that continuously regenerate themselves through their interactions and transformations, producing their own constituent states through networks of relational processes to form locally concrete spatial units, always struggling to maintain equilibrium (Varela, 1979).

In maintaining this precarious equilibrium, the organism uses its interactions with the environment, guiding itself in a way that allows it to revisit homeostatically favorable states over time, even in the face of continuous and random environmental fluctuations. And in doing so, it modifies its environment and its relationship to it, which also modifies the information that the organism will subsequently capture (Friston, 2005; Friston, 2010). The environment inhabited by an organism would therefore be nothing more than the consequences of its actions as captured by its sensory inputs, which strictly depend on what is afforded by a habitable context. Thus, we could not speak of an environment without there being an organism acting within it, as living beings bring forth the environment in which they inhabit/exist (Thompson, 1987). The distinction between an organism and its environment is consequently co-determinant, with the existence of both (as delimited entities) depending on their cyclic and dynamic interrelationships, over multiple nested scales (e.g. cell, tissue, organ, nestmate, colony, population, and ecosystem).

Considered on any scale (e.g. intelligence of nestmates and/or colonies [Feinerman; Korman, 2017]), the action of a thing in its environment depends first and foremost on its morphological and phenotypic constraints, creating inter-relational regularities that attract the organism to a specific set of conditions in which it may find itself throughout its life (Juarrero, 2015; Ramstead *et al.*, 2023). These inter-relational regularities that an organism tends to frequent throughout its existence can then be understood as a model of its niche (that is, the part of the environment that is accessible to it) and its behavior in it. And because such models establish something like a probabilistic mapping over time of the indirect interaction between the internal and external states of a living being, they can be interpreted as sub-personal (or sub-colony/nestmate, as it were) estimates of things about the environmental causes of changes in their constituent states. Such mapping would allow investigators to understand the internal dynamics of living beings as performing almost-optimal inferences about their niches (Corcoran; Pezzulo; Hohwy, 2020). The fundamental observation here is that this inferential process would provide organisms' behavior with a predictive nature. This allows the organism to respond to potential perturbations from their environment or to avoid physiologically stressful and homeostatically destabilizing conditions before they even occur – that is, allostasis. Corcoran *et al.* (2020) further specify that cognitive entities can accomplish allostasis by being able to “detach from the present and engage in counterfactual (active) inference”.

2.1 Active Inference

Building on the above considerations about agent-niche co-determination, Here we review various works that have analyzed

cognition and consciousness based on the Active Inference and the Free Energy Principle (FEP) (Ramstead *et al.*, 2023; Sandved-Smith *et al.*, 2021; Albarracín *et al.*, 2022), an approach which in our view integrates many of the valid and important points of the sections above.

So far, we have described an organism or thing that, in order to survive, needs to minimize entropy within its environment to distinguish itself from it and acquire autonomy over it. To maintain its distinction from the environment, the organism needs to act upon it, and by doing so repeatedly, it maps the regularities of the environment, preferring those that are favorable to its existence. In this context, its actions will increasingly be optimized concerning this favorable environment, enabling the organism to infer the causes of the effects that the environment has generated in its sensory apparatus. This is the starting point of the description provided by Active Inference, a recent neurocomputational framework that is based on a single organizing process: the minimization of prediction error and its role in the action-perception cycle.

According to the FEP, systems that minimize the divergence between their models and the environment information they sense will maintain their low (thermodynamic) entropy structure through autopoiesis in the face of random fluctuations in the environment. Under the FEP, phenotypes are cast as dynamics of systems (Friston *et al.*, 2023a; Friston *et al.*, 2023b), meaning that phenotypes are at least persistent enough to be repeatedly measured (e.g. in the scale-free quantum interpretation of FEP [Ramstead *et al.*, 2023; Fields; Friston; Glazebook; Levin, 2021]), if not self-evidencing enough to drive adaptive action in support of homeostasis. It is worth noting that all such modeling efforts can be considered as subjectively constructed

relational models (e.g., cognitive models created by a given modeler [Ramstead; Sakthivadivel; Friston, 2022; Parr; Pezzulo; Friston, 2022]).

Active Inference is a framework that describes the Bayesian-optimized functioning of behavior and cognition in living organisms in terms of the FEP (Corcoran; Pezzulo; Hohwy, 2020). This description follows only the imperative of minimizing surprise or variational free energy in the sensory observations that these organisms constantly receive (Parr; Pezzulo; Friston, 2022). Here, a system that possesses a physical brain is algorithmically organized in hierarchical layers that communicate via top-down and bottom-up information flows, ensuring the system's maintenance over time. The upper layers of the hierarchy deal with slower and more detailed information, transmitting it in a downward flow to anticipate the causes of the effects captured by sensors. The lower layers, closer to sensors, deal with faster and less detailed information as it arrives at the sensory apparatus and transmits it in an upward flow (Ramstead *et al.*, 2023; Clark, 2015). Only information that diverges from anticipatory predictions is propagated onward, in an ascending manner (e.g. from sensory organs upwards). This helps to maintain our active and updated predictive inferences, minimizing divergences and bringing our anticipatory dynamics closer to the world they aim to anticipate (Rauss; Schwartz; Pourtois, 2011, p. 1249). This minimization can occur in two ways: by updating organismic predictions (perception and learning) or by manipulating the environment to bring it closer to what is predicted by the system (action). In humans, the first form refers to the direction of attention, while the second form refers to embodied action upon the environment, altering it so that future sensory observations are more aligned with what the organism prefers (Clark, 2013, p. 186).

Thus, as applied to an empirical setting like an ant colony, the Active Inference framework consists of structuring relational, historically determined, and dynamic processes (Friedman *et al.*, 2021). In this context, systems are limited to capturing the sensory effects of causes in the world and depend on occasional regularities in these captures for the establishment of priors over time. Nonetheless, we need to act skillfully, even in the face of uncertainty in the environment, as we operate in the world to perceive it to the same extent that we perceive it to act in it. Therefore, in (nested) action-perception cycles, perceptual content informs the organism of what would be interesting to attend to and useful to interact with (i.e., its affordances) so that such a process can continue (Seth, 2014; Andersen; Miller; Vervaeke, 2022, p. 9).

3. Towards Physical and Metaphysical Dimensions for Cognitive Systems

Along with many previous works in this area our aims here are to devise and apply models and frameworks that support inquiries into consciousness, using assumptions that are more encompassing and at the same time operationalizable. Here we base and begin our contribution on the work of Birch *et al.* (2020) “Dimensions of Animal Consciousness”, where they proposed a multidimensional model to capture the variations of animal consciousness (Table 1). We do not claim that their model as presented is the total, complete, or only partitioning of the concepts around describing consciousness. Rather we see their proposal as, along with many other models in the space, a valid and relevant contender which merits consideration and juxtaposition, as well as further development and use. These kinds of pluralistic phenotypic approaches prevent collapsing down to any

unimodal scale (measurement), let alone one that centers around or calibrates to any specific cognitive system.

We pick up here on some of the closing words of Birch *et al.* (2020), who write: “We are not yet in a position to construct numerical measures of all five dimensions.” Here, we constructively attempt to pick up on that challenge, through the application of Active Inference to blend the empirical study of (eu)social complexity in insects, with a philosophical discussion of *What it is like to be an ant?*

3.1 Active Inference, Consciousness, and Social Complexity in Ants

In the preceding sections, we have laid the foundation to approach the quantitative multidimensional study of consciousness in biological entities as disparate as humans and ants, reviewing many current models and considering some aspects of the connection between consciousness and physical principles. By utilizing Active Inference, we are able to adopt a phenotype-oriented approach toward measuring consciousness, explored here in the context of ant colonies as tractable and global model study systems (Friedman; Søvik, 2019). As Table 1 explores, Active Inference plays an intermediary, or mapping, role between aspects of consciousness (on the left) and dimensions of social complexity phenotypes (on the right).

Area of Inquiry	Consciousness	Active Inference	Social complexity
Role played here	Operationalizing the multidimensional consciousness model of (Birch <i>et al.</i> 2020).	Bringing two measurement spaces together to support integrated inference.	Describes phenotypic aspects of insect colonies.
Epistemic status	Phenomenological, Axiomatic, High Road.	Frame & Filter – Grippled & Grippled.	Cognitive, Empirical Material, Low Road.

Temporality	"Integration across time". "a temporally integrated stream, rather than a staccato series of fragmented experiences".	Temporal depth or thickness. Time-cone/diamond historicity.	Variations in Life histories among species, colonies, and nestmates.
Unity	"integration at a time" "a single, unified perspective as opposed to multiple perspectives".	General synchrony. Mutual information. Time-cone/diamond width Deepest finite quantum holograph	Functional and material aspects of the ant colony, in terms of coherence, or causal / spatio-temporal extent of thing-ness
Selfhood (self-awareness)	"the conscious awareness of oneself as distinct from the world outside"	"I am therefore I think " (Friston, 2016)	Regulative or cybernetic features of the nested colony cognitive niche.
P(erceptual) richness	"the level of detail with which [animals] consciously perceive aspects of their environment. "	Sensory states Perception Sensitivity Salience	Colony and nestmate perceptual capacities, directly and through interactions and extended cognition.
E(valuative) richness	"the richness of [animals'] experiences of valence"	Valence Affect Basal and Higher-order (meta-) cognitive processes	Colony descriptive and generally regulative features, perhaps like level of alarm or overall activity patterns.

Table 1. Juxtaposition of Consciousness, Active Inference, and Social complexity. Note that one could add rows for (any number of) other cognitive phenomena and descriptors. Along the 3rd axis (normal to the page) can be the system of interest, here fixed as the ant colony so that this table can be printed flat on the classical plane.

Several points can be made about Table 1, considered here:

- As both consciousness and social complexity are potentially scale-independent descriptors for complex adaptive systems, their scope and study systems (e.g., humans, ant colonies, computers) are largely overlapping. Active Inference provides an approach and framework that can operationalize consciousness and social complexity studies, separately and jointly.
- Both consciousness and social complexity can be considered in terms of composite or complex variables, composed of various phenotypic sub-measures. A multidimensional method applied to study consciousness or social complexity does not have a simple “higher/lower” or “more/less” unless it is collapsed or projected down to a single comparable numerical value. Discussing insect evolution, Linksvayer and Johnson (2019), challenged the concept of a singular hierarchical social ladder structure for social complexity. Their critique highlights how numerical measurements for social complexity are often discussed in a rankable fashion (e.g., more/less socially complex, or higher/lower in terms of social complexity). However, distinct insect lineages show diverse changes across phenotypic measures of social complexity over multiple phylogenetic scales. Consequently, they argue against oversimplifying the range of ant colony phenotypes into a linear or ladder-like measure, arguing that multidimensional measures will be useful as more information on insect sociality is analyzed through time. We feel a strong analogical resonance between their discussion of social complexity in ants, and consciousness studies in ants and beyond.

- Empirically, different axes of social complexity are not necessarily correlated with each other (e.g., morphologically distinct workers may be found in large or small colonies) (Holland; Bloch, 2020), just as different proposed measures for cognitive function or consciousness in humans do not always correlate. As a corollary to this point, empirical cognitive measures (e.g. neuron number or connectivity in brains, nestmate number or connectivity in colonies) may have complex or unclear correlations to any dimensions of consciousness.
- Certain measures of social complexity in ants are similar or even identical to specific aspects of consciousness, for example in the domains of information processing capacity, temporal depth, and the ability to execute advanced distributed strategies. These kinds of network information-processing properties are key aspects of the IIT and other neurocognitive consciousness theories, yet bear little metaphysical baggage/interpretation in ants – suggesting that mere and even total empirical measurement, may not resolve philosophical debates.

4. Discussion

Here we have reviewed, echoed, and extended the calls of others to study consciousness in terms of multi-dimensional spaces. This perspective not only paves the way for a more comprehensive understanding of consciousness but also has the potential to impact various fields, such as medicine, psychology, ethics, artificial intelligence, and our daily lives. We proposed here that consciousness

studies should proceed with the understanding that consciousness shares many properties, and even overlapping referents, with social complexity in ants.

While it is clear that there is no linear hierarchy, or “Scala Naturae”, for insect sociality (Linksvayer; Johnson, 2019), the notion of a “Scala Conscientia” — a singular, ascending scale of consciousness in which “more” and “less” consciousness could exist — within or across species remains implicitly or explicitly entrenched. The philosophical and conceptual analysis provided here supports an open and rigorous perspective which we believe will be of great epistemic relevance and pragmatic utility for the field of consciousness studies.

Watching ants and speculating about their consciousness and cognition is probably a past-time as old as time (Proverbs, 2018). Here we’ve argued that Active Inference today can play a catalytic and integrative role that bridges the study of consciousness with social complexity. As behavioral myrmecology is not traditionally looking for consciousness as part of the causal account of a colony, the field offers a variety of systems (ants), (extended cognitive) theories, and diverse measures that consciousness researchers and philosophers may delight in alike for years to come.

Open & Continuing Questions:

- How can we make comparisons among different complex systems, in terms of different criteria for consciousness and cognition?
- How can existing models of consciousness be adapted or integrated within broader multidimensional models?
- How do measures of consciousness in ant colonies relate to measures of social complexity?

- How can recent advances in multiscale systems help us create models of consciousness in diverse cognitive systems?
- What is the role of the action-perception interface and physics of information processing in consciousness, and how do different types of phenotypic measures help us understand this?

References

- ALBARRACIN, M. *et al.* Mapping Husserlian phenomenology onto active inference. *arXiv [q-bio.NC]*, 2022. Available at: <http://arxiv.org/abs/2208.09058>.
- ANDERSEN, B. P.; MILLER, M.; VERVAEKE, J. Predictive processing and relevance realization: exploring convergent solutions to the frame problem. *Phenomenology and the Cognitive Sciences*, 2022. DOI: 10.1007/s11097-022-09850-6
- ANDERSON, C.; MCSHEA, D. W. Individual versus social complexity, with particular reference to ant colonies. *Biological Reviews of the Cambridge Philosophical Society*, v. 76, p. 211-237, 2001. DOI: 10.1017/s1464793101005656.
- ANDREWS, M. The math is not the territory: Navigating the free energy principle. *Biology & Philosophy*, v. 36, p. 30, 2021. DOI: 10.1007/s10539-021-09807-0.
- ATMANSPACHER, H. Quantum approaches to consciousness. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2020. Available at: <https://plato.stanford.edu/archives/sum2020/entries/qt-consciousness/>.
- BAARS, B. J. *A cognitive theory of consciousness*. Cambridge University Press, 1988. Available at: <https://philpapers.org/rec/BAAACT>.
- BAARS, B. J. The global workspace theory of consciousness. *The Blackwell companion to consciousness*, Chichester, UK: John Wiley & Sons, Ltd, 2017. p. 227-242. DOI: 10.1002/9781119132363.ch16.
- BARÚSS, I. Metanalysis of definitions of consciousness. *Imagination, Cognition and Personality*, v. 6, p. 321-329, 1987. DOI: 10.2190/39X2-HMUL-WB7B-B1A1.19.

- BAYN, T.; HOHWY, J.; OWEN, A. M. Are there levels of consciousness? *Trends in Cognitive Sciences*, v. 20, p. 405-413, 2016. DOI: 10.1016/j.tics.2016.03.009.
- BIRCH, J.; SCHNELL, A. K.; CLAYTON, N. S. Dimensions of animal consciousness. *Trends in Cognitive Sciences*, v. 24, p. 789-801, 2020. DOI: 10.1016/j.tics.2020.07.007.
- BOOLS, G. On “seeing” the truth of the Gödel sentence. *Behavioral and Brain Sciences*, v. 13, p. 655-656, 1990. DOI: 10.1017/S0140525X00080687
- BONGARD, J.; LEVIN, M. There’s plenty of room right here: Biological systems as evolved, overloaded, multi-scale machines. *Biomimetics*, v. 8, 2023. DOI: 10.3390/biomimetics8010110.
- BRESSLER, S. L.; KELSO, J. A. S. Coordination dynamics in cognitive neuroscience. *Frontiers in Neuroscience*, v. 10, n. 397, 2016. DOI: 10.3389/fnins.2016.00397.
- CIAUNICA, A. *et al.* The first prior: From co-embodiment to co-homeostasis in early life. *Consciousness and Cognition*, v. 91, p. 103117, 2021. DOI: 10.1016/j.concog.2021.103117.
- CLARK, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, v. 36, p. 181-204, 2013. DOI: 10.1017/S0140525X12000477.
- CLARK, A. *Surfing uncertainty: Prediction, action, and the embodied mind.* Oxford University Press, 2015. DOI: 10.1093/acprof:oso/9780190217013.001.0001
- CORCORAN, A. W.; PEZZULO, G.; HOHWY, J. From allostatic agents to counterfactual cognisers: Active inference, biological regulation, and the origins of cognition. *Biology and Philosophy*, v. 35, p. 32, 2020. DOI: 10.1007/s10539-020-09746-2.
- CORDES, R. J.; FRIEDMAN, D. A.; MARON, M. *Reimagining maps.* 2020. DOI: 10.5281/zenodo.4170026
- COSTA, J. T. Scale models? What insect societies teach us about ourselves. *Proceedings of the American Philosophical Society*, v. 146, p. 170-180, 2002.
- CRICK, F.; KOCH, C. Toward a neurobiological theory of consciousness. *Seminars in Neuroscience*, v. 2, p. 263-275, 1990. Available at: <https://philpapers.org/rec/CRITAN>.

- DEL PIN, S. H. *et al.* Comparing theories of consciousness: Why it matters and how to do it. *Neuroscience of Consciousness*, v. 2021, niab019, 2021. DOI: 10.1093/nc/niab019.
- DEHAENE, S.; CHANGEUX, J.-P. Experimental and theoretical approaches to conscious processing. *Neuron*, v. 70, p. 200-227, 2011. DOI: 10.1016/j.neuron.2011.03.018.
- DENNETT, D. C. *Consciousness explained*. Nova York, NY: Little, Brown & Company, 1992.
- DERAKHSHANI, M. *et al.* At the crossroad of the search for spontaneous radiation and the Orch OR consciousness theory. *Physical Life Review*, v. 42, p. 8-14, 2022. DOI: 10.1016/j.plrev.2022.05.004.
- FEINERMAN, O.; KORMAN, A. Individual versus collective cognition in social insects. *Journal of Experimental Biology*, v. 220, p. 73-82, 2017. DOI: 10.1242/jeb.143891
- FIELDS, C.; FRISTON, K.; GLAZEBOOK, J. F.; LEVIN, M. A free energy principle for generic quantum systems. *arXiv [quant-ph]*, 2021. Available at: <http://arxiv.org/abs/2112.15242>
- FRIEDMAN, D. A.; SØVIK, E. The ant colony as a test for scientific theories of consciousness. *Synthese*, p. 1-24, 2019. DOI: 10.1007/s11229-019-02130-y.
- FRIEDMAN, D. A. *et al.* Active inferants: The basis for an active inference framework for ant colony behavior. *Frontiers in Behavioral Neuroscience*, v. 15, p. 126, 2021. DOI: 10.3389/fnbeh.2021.647732.
- FRISTON, K. A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, v. 360, p. 815-836, 2005. DOI: 10.1098/rstb.2005.1622.
- FRISTON, K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci.* 2010;11: 127-138. DOI:10.1038/nrn2787
- FRISTON, K. I am therefore I think. *The unconscious: A bridge between psychoanalysis and cognitive neuroscience.* 2016; 127-151. DOI:10.4324/9781315686998-17
- FRISTON, K. *et al.* Path integrals, particular kinds, and strange things. *Physical Life Review*, v. 47, p. 35-62, 2023a. DOI: 10.1016/j.plrev.2023.08.016.

- FRISTON, K. J. *et al.* A variational synthesis of evolutionary and developmental dynamics. *Entropy*, v. 25, 2023b. DOI: 10.3390/e25070964
- GORDON, D. M. Wittgenstein and ant-watching. *Biology and Philosophy*, v. 7, p. 13-25, 1992. DOI: 10.1007/BF00130161.
- GORDON, D. M. *Ant encounters: Interaction networks and colony behavior*. Princeton University Press, 2010.
- GRAZIANO, M. S. A. *et al.* Toward a standard model of consciousness: Reconciling the attention schema, global workspace, higher-order thought, and illusionist theories. *Cognitive Neuropsychology*, v. 37, p. 155-172, 2020. DOI: 10.1080/02643294.2019.1670630.
- GRUSH, R.; CHURCHLAND, P. S. Gaps in Penrose's Toiling. *Journal of Consciousness Studies*, v. 2, p. 10-29, 1995. Available at: <https://philpapers.org/rec/CHUGIP>.
- HAMEROFF, S.; PENROSE, R. Consciousness in the universe: A review of the "Orch OR" theory. *Physical Life Review*, v. 11, p. 39-78, 2014. DOI: 10.1016/j.plrev.2013.08.002.
- HAMEROFF, S. "Orch OR" is the most complete, and most easily falsifiable theory of consciousness. *Cognitive Neuroscience*, p. 74-76, 2021. DOI: 10.1080/17588928.2020.1839037.
- HEYLIGHEN, F. Stigmergy as a universal coordination mechanism I: Definition and components. *Cognitive Systems Research*, v. 38, p. 4-13, 2016. DOI: 10.1016/j.cogsys.2015.12.002.
- HOFSTADTER, D. R. *Godel, Escher, Bach: An eternal golden braid*. 2 Ed. Basic Books, 1999.
- HOLLAND, J. G.; BLOCH, G. The complexity of social complexity: A quantitative multidimensional approach for studies of social organization. *The American Naturalist*, v. 196, p. 525-540, 2020. DOI: 10.1086/710957.
- JI, X. *et al.* Sources of richness and ineffability for phenomenally conscious states. *arXiv [q-bio.NC]*, 2023. Available at: <http://arxiv.org/abs/2302.06403>.
- JUARRERO, A. What does the closure of context-sensitive constraints mean for determinism, autonomy, self-determination, and agency? *Progress in Biophysics and Molecular Biology*, v. 119, p. 510-521, 2015. DOI: 10.1016/j.pbiomolbio.2015.08.007.

- KASS, J. M. *et al.* The global distribution of known and undiscovered ant biodiversity. *Science Advances*, v. 8, eabp9908, 2022. DOI: 10.1126/sciadv.abp9908.
- KING, S. *Insect nations: Visions of the ant world from Kropotkin to Bergson*. InkerMen Press, 2006.
- KOCH, C.; HEPP, K. Quantum mechanics in the brain. *Nature publishing group UK* [Internet]. 29 mar. 2006. DOI: 10.1038/440611a
- LEDoux, J. E.; MICHEL, M.; LAU, H. A little history goes a long way toward understanding why we study consciousness the way we do today. *Proceedings of the National Academy of Sciences of the United States of America*, v. 117, p. 6976-6984, 2020. DOI: 10.1073/pnas.1921623117.
- LINKSVAYER, T. A. Chapter Eight - The molecular and evolutionary genetic implications of being truly social for the social insects. In: ZAYED, A.; KENT, C. F. (Eds.). *Advances in insect physiology*. Academic Press, p. 271-292, 2015. DOI: 10.1016/bs.aiip.2014.12.003.
- LINKSVAYER, T. A.; JOHNSON, B. R. Re-thinking the social ladder approach for elucidating the evolution and molecular basis of insect societies. *Current Opinion in Insect Science*, v. 34, p. 123-129, 2019. DOI: 10.1016/j.cois.2019.07.003
- LONGINO, H. E. Toward an epistemology for biological pluralism. In: *Biology and epistemology*, p. 262-286, 2000.
- MULLEN, R. J.; MONEKOSSO, D.; BARMAN, S.; REMAGNINO, P. A review of ant algorithms. *Expert Systems with Applications*, v. 36, p. 9608-9617, 2009. DOI: 10.1016/j.eswa.2009.01.020.
- NAGEL, T. *The view from nowhere*. Oxford University Press, 1989.
- NEGRO, N. Emergentist integrated information theory. *Erkenntnis*, 2022. DOI: 10.1007/s10670-022-00612-z.
- ORORBIA, A.; FRISTON, K. Mortal computation: A foundation for biomimetic intelligence. *arXiv [q-bio.NC]*, 2023. Available at: <http://arxiv.org/abs/2311.09589>
- PARR, T.; PEZZULO, G.; FRISTON, K. J. *Active inference: The free energy principle in mind, brain, and behavior*. MIT Press, 2022.

- PUTNAM, H. Review of emphshadows of the mind. *AMS Bulletin*. 1995; 32. Available at: https://philpapers.org/rec/PUTRO_
- PROVERBS. *Bible* [online]. Available at: <https://www.biblegateway.com/passage/?search=Proverbs+6%3A6-8&version=NIV>.
- RAMSTEAD, M. J. D.; BADCOCK, P. B.; FRISTON, K. J. Answering Schrödinger's question: A free-energy formulation. *Physical Life Review*, v. 24, p. 1-16, 2018. DOI: 10.1016/j.pprev.2017.09.001.
- RAMSTEAD, M. J. D.; SAKTHIVADIVEL, D. A. R.; FRISTON, K. J. On the map-territory fallacy fallacy. *arXiv [physics.hist-ph]*, 2022. Available at: <http://arxiv.org/abs/2208.06924>
- RAMSTEAD, M. J. D. *et al.* On Bayesian mechanics: A physics of and by beliefs. *Interface Focus*, v. 13, p. 20220029, 2023. DOI: 10.1098/rsfs.2022.0029.
- RAMSTEAD, M. J. D. *et al.* The inner screen model of consciousness: Applying the free energy principle directly to the study of conscious experience. *PsyArXiv*, 2023. DOI: 10.31234/osf.io/6af3
- RAUSS, K.; SCHWARTZ, S.; POURTOIS, G. Top-down effects on early visual processing in humans: A predictive coding framework. *Neuroscience & Biobehavioral Reviews*, v. 35, p. 1237-1253, 2011. DOI: 10.1016/j.neubiorev.2010.12.011.
- ROROT, W. Bayesian theories of consciousness: A review in search for a minimal unifying model. *Neuroscience of Consciousness*, v. 2021, p. niab038, 2021. DOI: 10.1093/nc/niab038
- ROSENTHAL, D. Concepts and definitions of consciousness. *Encyclopedia of consciousness*, p. 157-169, 2009.
- SAFRON, A. Integrated World Modelling Theory (IWMT) implemented: Towards reverse engineering consciousness with the free energy principle and active inference. *Active Inference*. Springer International Publishing, 2020. p. 135-155. DOI: 10.1007/978-3-030-64919-7_15

SANDVED-SMITH, L. *et al.* Towards a computational phenomenology of mental action: Modelling meta-awareness and attentional control with deep parametric active inference. *Neuroscience of Consciousness*, v. 2021, p. niab018, 2021. DOI: 10.1093/nc/niab018.

SAPOLSKY, R. M. *Determined: The science of life without free will.* 2023.

SETH, AK. The Cybernetic Bayesian Brain. *Open MIND.* Johannes Gutenberg Universität Mainz; 2014. DOI:10.15502/9783958570108

SETH, A. *Being you: A new science of consciousness.* Penguin, 2021.

SETH, AK.; BAARS, B. J.; EDELMAN, D. B. Criteria for consciousness in humans and other mammals. *Consciousness and Cognition*, v. 14, p. 119-139, 2005. DOI: 10.1016/j.concog.2004.08.006.

SKINNER, B. F. *Beyond freedom and dignity.* Hackett Publishing Company, Inc., 1971.

STORM, J. F. *et al.* An integrative, multiscale view on consciousness theories. Early version of paper in press in *NEURON*, 2024. DOI: 10.31234/osf.io/9byzu

SCHRÖDINGER, E. *What is life? With mind and matter and autobiographical sketches.* Cambridge University Press, 1992.

STEWART J, GAPENNE O, DI PAOLO, EA. *Enaction: Toward a New Paradigm for Cognitive Science.* MIT Press; 2010.

TASSI, P.; MUZET, A. Defining the states of consciousness. *Neuroscience & Biobehavioral Reviews*, v. 25, p. 175-191, 2001. DOI: 10.1016/s0149-7634(01)00006-9.

TEGMARK, M. The importance of quantum decoherence in brain processes. *arXiv [quant-ph]*, 1999. Available at: <http://arxiv.org/abs/quant-ph/9907009>.

THOMPSON, WI. *Gaia, a Way of Knowing: Political Implications of the New Biology.* *Inner Traditions International*; 1987. Available at: <https://philpapers.org/rec/THOGAW-2>

TINDALL, S. C. *Level of consciousness.* Butterworths, 1990. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK380/>.

TONONI, G. An information integration theory of consciousness. *BMC Neuroscience*, v. 5, p. 42, 2004. DOI: 10.1186/1471-2202-5-42.

TRANIELLO, J. F.; LINKSVAYER, T. A.; COTO, Z. N. Social complexity and brain evolution: Insights from ant neuroarchitecture and genomics. *Current Opinion in Insect Science*, v. 53, p. 100962, 2022. DOI: 10.1016/j.cois.2022.100962

TSUCHIYA, N.; TAGUCHI, S.; SAIGO, H. Using category theory to assess the relationship between consciousness and integrated information theory. *Neuroscience Research*, v. 107, p. 1-7, 2016. DOI: 10.1016/j.neures.2015.12.007

VARELA, FJ. *Principles of Biological Autonomy*. North-Holland; 1979. Available at: <https://philpapers.org/rec/VARPOB>

VELMANS, M. How to define consciousness - And how not to define consciousness. *Journal of Consciousness Studies*, v. 16, p. 139-156, 2009. Available at: <http://research.gold.ac.uk/26130/>.

4

A MEMÓRIA IMUNOLÓGICA COMPOSTA POR FORMIGAS

Aryel C. Goes

“A característica mais comum de um *organismo* em uma colônia de formigas é a sua individualidade. Assim como a célula ou uma pessoa, a colônia se comporta como uma *unidade individual*, mantendo sua identidade no espaço, resistindo a dissoluções e, como regra geral, qualquer tipo de fusão com colônias de mesma ou diferente espécie”.

(William Morton Wheeler, 1911, tradução livre)

“Parece provável que, em alguns casos, o *sistema imune* de uma colônia a tornará fortemente coesiva de forma que isso a qualificará como um “*organismo*”. Além disso, esses casos podem evidenciar a transição da multicelularidade e de transições evolutivas”.

(Thomas Pradeu, 2012, p. 12, tradução livre)

Ao longo deste livro, aspectos filosóficos e metafísicos lhe foram apresentados para a sugestividade de uma consciência externalizada do cérebro, formada por várias unidades individuais de insetos vivendo em sociedade. À parte, tal sugestiva consciência também precisa ser elucidada na perspectiva *biológica*. Para tanto, muitos exemplos declarados e estudados pela professora Deborah M. Gordon (citados ao longo do livro) apresentam a integração de experiências individuais para compor o comportamento geral em colônias de formigas (Gordon, 2010). Muitos desses exemplos são pautados em forrageio, fundação de ninhos, resposta à predações e competição por recursos, alteração no comportamento pelas mudanças ambientais, e dentre outros. Embora pouco citada, o campo da *imunologia* também observa a relevância das experiências individuais na composição da resposta imune de uma colônia contra ameaças externas.

Pautada em evidências biológicas atuais, a consciência externalizada da colônia também pode ser investigada pela perspectiva das respostas higiênicas em insetos sociais, conhecida como *imunologia social* (Cremer *et al.*, 2007). Essa é a conhecida imunologia do *superorganismo*, ou melhor definida como “o sistema imune da colônia”. Digo isso pela forma com que espécies de formigas e, outros insetos sociais, são capazes de modificar o seu padrão de resposta higiênica devido à experiência prévia contra patógenos. Em resultado, a plasticidade desse comportamento sanitário pode tornar as medidas protetoras e defensivas mais eficientes, aumentando a taxa de sobrevivência dos indivíduos que compõem a colônia. Comparando o nosso sistema imunológico com o de uma colônia de formigas, seria o equivalente à imunização auxiliar o sistema imune em respostas mais rápidas e específicas em futuras infecções ao mesmo patógeno. Interessantemente, a isso podemos hipotetizar uma *memória imune coletiva* em formigas.

Antes de aprofundarmos a hipótese de memória imune coletiva e como ela poderia embasar a consciência externalizada se confirmada, precisamos entender a relevância da imunologia social para uma colônia de formiga. Primeiramente, vamos visualizar um pouco o papel do sistema imune para a *individualização* de um organismo (Pradeu, 2012); desta forma, teremos mais claro o motivo de podermos tratar uma colônia como unidade biológica íntegra, assim dita na epígrafe do texto de Wheeler (1911). Em seguida, serão apresentados os princípios da imunologia social, utilizando como organismo modelo as formigas cortadeiras de folha (e.g. saúvas e quenquéns). Com isso explicado, poderei discorrer sobre as sugestividades e indícios de que as respostas imunes coletivas de tais formigas (e outras espécies) são plásticas devido a experiências, baseada em estudos recentes. Assim, vou

comentar os aspectos ecológicas e evolutivas dessa suposta memória imune social e as suas implicações para a teoria da consciência expandida.

1. O sistema imune, o indivíduo e a quimera

Você consegue definir o que é *sistema imune*? Provavelmente, “defesa contra patógenos ou doenças” é a primeira definição que lhe vem à mente. De antemão não está errado, e é uma das terminologias que comumente encontrará em livros, sites e aprenderá no ensino básico e superior. Porém, tal definição está incompleta. Hoje, muitos biólogos e filósofos da ciência reconhecem e apresentam dados empíricos que suportam funções mais estritas para o sistema imunológico, ampliando, portanto, a sua definição.

1.1 Imunologia e individualidade

Além da conhecida função para combater doenças e substâncias estranhas, o sistema imune tem um papel crucial na *individualidade* de um organismo, já que ele é capaz de estabelecer limites ao mesmo (Pradeu, 2011, 2012). O argumento para isso vem do chamado *princípio da inclusão*, que estabelece o quê o corpo físico de um organismo irá rejeitar ou aceitar dentro de sua delimitação física (Pradeu, 2011, 2012). Assim, o sistema imune reconhece o quê faz parte do indivíduo e deve permanecer nele, e o quê deve ser rejeitado. Observamos parcialmente esse princípio quando notamos as dificuldades e cuidados que médicos precisam ter para realizar um transplante de tecido ou órgão, pela resposta autoime do paciente receptor. Os autores Gould e Lloyd (1999, p. 11906) deixam isso claro ao estabelecer que “organismos são coerentemente associados no espaço, mantendo-se reconhecíveis por

uma forma física, delimitada pela pele, que separa a si próprio do que é do mundo externo [...] o sistema imune é quem possui os aparatos para isso, reconhecendo e desarmando transgressores que adentram esse espaço interior”.

Essa frase de Goyld e Lloyd resume o famoso modelo da imunologia, o *self* e *nonself* (aquilo que é e não é do indivíduo). De acordo com essa teoria, tudo que seja estranho e vem de fora do organismo (*nonself*) incitará uma resposta imunológica, enquanto aquilo que faz parte do organismo e que não esteja se comportando da maneira ‘correta’ (*self*) também irá ativar resposta imune. Em exemplo, uma vírus ou bactéria pode ser considerado como algo estranho que vem de fora (*nonself*). Por outro lado, o câncer é composto de células que, embora façam parte do mesmo organismo, não estão se comportando corretamente (*self*) – ou com semelhança ao resto do padrão molecular.

Fora o sistema imune, há uma hipótese alternativa que auxiliaria na delimitação do organismo multicelular: a *homogeneidade genética* (Pradeu, 2012). Ela preza que constituintes do organismo que apresentam o mesmo material genético (são iguais, homogêneas) irão cooperar melhor do que se fossem heterogêneas (geneticamente distintas), já que possuem a *mesma função*. A separação de células somáticas (aquelas que compõem o tecido da pele) das células germinativas (as células sexuais; gametas) seria um exemplo deste tipo. Não há razões cabíveis para que células da epiderme queiram se diferenciar do todo e ter sua independência reprodutiva, aumentando o chamado *fitness* próprio. E mesmo que houvesse tal tentativa de corrupção, o sistema imune será capaz de combater tais células que queiram se tornar diferentes do todo, ou remover aquelas que sejam tão estranhas (i.e., antígenos, patógenos) (Buss, 1987; Michod, 1999). Novamente, temos o exemplo do câncer. Células cancerígenas ativam

respostas específicas do sistema imune as quais, na maioria dos casos, levam em sua destruição. Desta forma, a capacidade imunológica tem papel decisivo na estabilidade e manutenção do organismo.

1.2 A Quimera

Embora tais hipóteses sejam conhecidas e defendidas por muitos imunologistas e filósofos, a homogeneidade genética é aquela que precisa ser repensada. Por não ser uma condição encontrada na maioria dos seres vivos, essa não deve ser vista como um contribuidor da evolução do indivíduo (Michod, 1999; Strassmann, 2004). A razão para isso é que todo organismo multicelular é heterogêneo por si próprio (Pradeu, 2010). Invertebrados, vertebrados e plantas abrigam muitas bactérias e fungos comensais e simbióticas dentro de si (Garrett *et al.*, 2010; Triviño, Suárez, 2020). Portanto, qualquer organismo biológico é como uma *quimera*. Muitas bactérias simbiontes que vivem em nosso corpo, por exemplo, apresentam funções indispensáveis na digestão, imunologia e até no desenvolvimento (Pradeu, 2010 e referências associadas). Não estão presentes por mero acaso, mas sim por serem parte do que constitui o organismo como um todo, o conhecido: *holobionte* (Triviño, Suárez, 2020). Uma vez que o modelo *self/nonsel* apresentado acima prediz a radicalização pelo sistema imune de seres diferentes do próprio corpo (*nonsel*), por qual razão o ser humano ou outros animais conseguem viver em simbiose com sua microbiota? A mesma não passa despercebida pelo nosso corpo e ao sistema imune, do contrário: estão continuamente em interação, como hipotetizado pela *teoria da continuidade* de Pradeu (2012).

Agora, finalmente podemos chegar a uma boa conclusão e definição do que é o *sistema imune*, e prosseguirmos para o tópico

principal deste capítulo. O que incita a reação imune não é necessariamente algo geneticamente estranho ao corpo/organismo, mas sim àquilo que apresenta uma aparência incomum do “padrão molecular”. A isso designa-se a *teoria da continuidade*, sugerindo que qualquer resposta imunológica decorre da detecção de padrões moleculares que sejam diferentes daqueles que o sistema imune comumente interage; sejam eles de origem endógena (do próprio organismo, como células cancerígenas, microbiota, etc.) ou exógena (como bactérias e fungos patogênicos, vírus, etc) (Figura 1). Assim, um organismo multicelular é composto de constituintes *heterogêneos*, sendo parte desses originários desde o ‘nascimento’ (células), enquanto outros são obtidos ao longo da vida (i.g., bactérias simbiontes), ao qual a sua coesividade e integridade serão estabelecidas. E quem a estabelecerá? *O sistema imunológico*.

Concluindo, a imunidade desempenha um papel além da proteção contra doenças. “Ela é um mecanismo crucial na definição de um indivíduo e organismo biológico, seja por delimitar seus limites espaciais (*princípio da inclusão*) e policiar contra aquilo que seja anormal do padrão interagido (*hipótese da continuidade*)” (Pradeu, 2012). Veja que a definição generalista ‘proteção contra doenças’ que tanto disseminamos é apenas uma parcela das várias funções do sistema imune para manter a coesividade do organismo. Como ressaltado no início, o termo não está errado ou defasado, apenas *incompleto*.

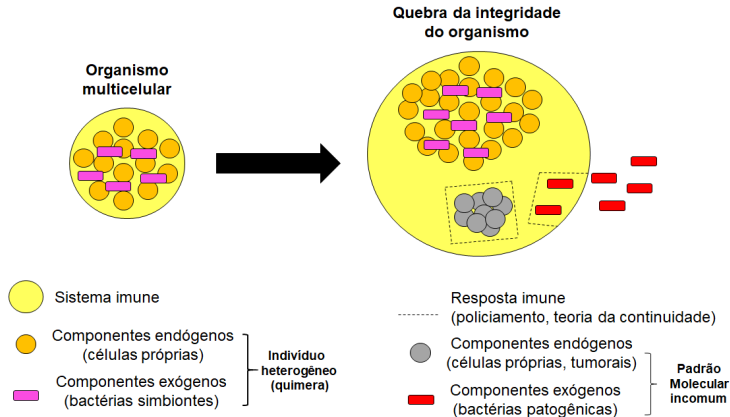


Figura 1. A funcionalidade do sistema imune está além de proteger o indivíduo contra doenças e patógenos. Hoje em dia há uma visão mais branda da sua importância para manter a **integridade do organismo**, seja para (i) combater potenciais “trapaceiros” endógenos que se diferenciam muito do todo (ex: células cancerígenas), ou (ii) componentes endógenos que sejam incomuns (ex: bactérias patogênicas). Ele mantém a delimitação física e coesiva entre os componentes do organismo. Desta forma, o organismo é heterogêneo, composto de si e de outros simbiotes.

2. O sistema imune da colônia

Agora que entendemos melhor a funcionalidade do sistema imune e a sua relevância para a integridade de um organismo, podemos passar adiante de sua relevância em organismos multicelulares para a de *superorganismos*. Se lembra da epígrafe de William Morton Wheeler? Pois bem, finalmente podemos aprofundá-la.

Um *superorganismo* é uma reunião de organismos similares que, em conjunto, podem ser considerados como uma única unidade individual. Isso ocorre por alguns aspectos básicos: (i) integração, (ii) coesividade, (iii) divisão de trabalho, e (iv) separação entre castas reprodutivas e não-reprodutivas (Boomsma e Gawne, 2018; Pull e McMahon, 2020). Quer um exemplo no reino animal? Wheeler nos concede: uma *colônia de formigas*. Hoje, aqueles que estudam sobre a evolução da socialidade, consideram as formigas como seres altamente socializados (se

comparados com outros: vespas e cupins eusociais), encaixando perfeitamente nos quatro pressupostos acima citados, podendo então serem chamadas por superorganismo; algo que em sua natureza e conformação age como uma única coisa: uma entidade individual. Assim como um organismo multicelular, a colônia de formigas também precisa manter sua integridade – seja contra seres externos (i.g., patógenos ou invasores de outras espécies) ou aqueles que compõem a própria colônia (i.g., operárias irmãs). Considerando isso, aonde se encaixa o sistema imune?

Insetos, no geral, apresentam um sistema imune. Obviamente que cada operária da colônia tem o seu próprio sistema imune celular. Além dela, sabe-se da existência do sistema imune expressado em um nível superior, além do fisiológico das formigas. Este não é composto por células, mas sim pelas próprias operárias (Cremer *et al.*, 2007; Cremer, Sixt, 2009; Pull, McMahon, 2020). *Imunologia social* é o termo dado por Cremer e colaboradores (2007) para o conjunto de mecanismos fisiológicos, comportamentais e organizacionais empregados em conjunto ou, individualmente, pelas operárias que compõem uma colônia. Da mesma forma que comentei acima, esse sistema imune social têm funções relevantes para manter a colônia íntegra e individual, assim como bem exemplificado pela fala de Wheeler (1911). Quais seriam esses mecanismos? Como essa imunologia *sem células* é expressada pela colônia?

2.1 A imunologia social em formigas cortadeiras

Assim como em qualquer organismo, vírus, protozoários, fungos e bactérias também são facilmente transmitidos entre companheiras de ninho e proles da colônia, afetando o fitness da mesma (Schmid-

Hempel, 1998; Cremer *et al.*, 2007; Meunier, 2015). Embora a imunologia social tenha evoluído para proteger os indivíduos do ninho que compõem o superorganismo, alguns insetos sociais ao longo de sua história evolutiva adquiriram parceiros simbióticos que também necessitam de igual proteção (Currie *et al.*, 1999). As formigas e cupins que *cultivam fungos* representam um exemplo notável desse desafio. Seu envolvimento no cultivo de fungos como principal fonte nutricional permitiu que esses insetos dominassem o Novo e o Velho Mundo (Möller, 1893; Weber, 1972; Mueller, 2002; Aanen *et al.*, 2002), respectivamente, explorando novos nichos ecológicos. Como desvantagem, eles possuem a tarefa adicional de proteger o companheiro fúngico de microrganismos externos que podem ameaçar o seu mutualismo (Currie e Stuart, 2001; Otani *et al.*, 2019; Goes *et al.*, 2020).

Além de enfrentarem os mesmos obstáculos vistos em outros insetos sociais, as formigas e cupins cultivadores de fungos (Attina e Macrotermitidae, respectivamente) também precisam proteger o seu mutualista. A degradação do substrato vegetal utilizado para adubar o jardim de fungo o torna extremamente rico em nutrientes compostos por polissacarídeos e açúcares simples (Martin, 1969; Siqueira *et al.*, 1998; Silva *et al.*, 2006a). Apesar de serem importantes para a nutrição das operárias e da ninhada, tais nutrientes oferecem oportunidades para vários microrganismos explorarem a colônia. No caso das formigas cortadeiras (gêneros *Acromyrmex* e *Atta*), muitos fungos externos podem ser patógenos generalistas, especializados (Currie *et al.*, 1999; Currie; Stuart, 2001; Rodrigues *et al.*, 2008, 2011), ou endofíticos (fungos que vivem dentro do tecido vegetal, Van Bael *et al.*, 2009). Independentemente, muitos espécimes são conhecidos pela sua capacidade em afetar negativamente o fungo das formigas. Assim, a

domesticação de *Leucoagaricus gongylophorus* (Agaricales, Agaricaceae) requereu adaptações adicionais da imunologia social para manter a fungicultura viva (Figura 2).

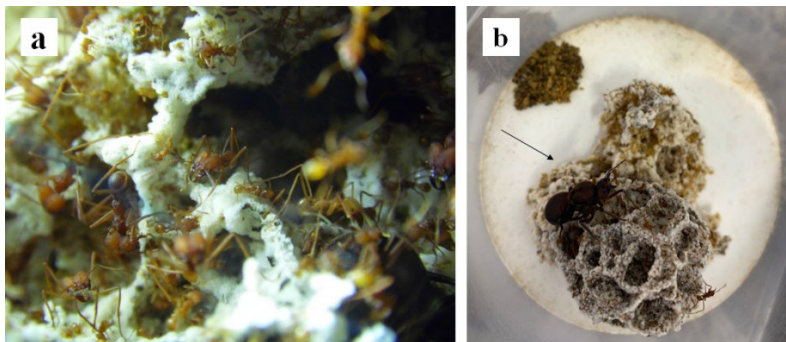


Figura 2. A formiga cortadeira de folha, *Atta sexdens* (saúva-limão) e o seu fungo basidiomiceto, *Leucoagaricus gongylophorus*. a) Operárias de *A. sexden* junto ao *L. gongylophorus* em sua estrutura de cultivo conhecida por jardim de fungo. b) Colônia incipiente (jovem) com aproximadamente 3 meses de idade; note a presença da rainha no topo do jardim de fungo (indicado pela seta), junto com algumas operárias e o descarte de lixo ao lado. Fotos: o autor.

Em colônias de formigas comuns, a imunidade social é direcionada aos indivíduos em si, com ações individuais e grupais que protegem toda a colônia (Cremer *et al.*, 2007). No entanto, para as formigas que cultivam fungo tais mecanismos também precisam ser aplicados para a proteção do seu fungo mutualista. Em formigas cortadeiras, as operárias podem aplicar estratégias químicas ou físicas para prevenir ou conter infecções. Formigas secretam e aplicam compostos de suas glândulas metapleurais (com atividade fungistática, fungicida e bacteriostática) em seu corpo para prevenir o crescimento de microrganismos entomopatogênicos (Hölldobler, 1984; Graystock, Hughes, 2011). Além disso, essa ação também é direcionada à superfície do jardim do fungo, garantindo o não desenvolvimento de fungos antagonistas ou maléficos para o mesmo (Ortius-Lechner *et al.*, 2000;

Bot *et al.*, 2002; Poulsen *et al.*, 2002; Fernández-Marín *et al.*, 2006; 2015). Outro exemplo é com o fungo de gênero *Escovopsis*, considerado prejudicial ao jardim de fungo. Ele pode ser inibido por compostos antibióticos produzidos por tais glândulas (Currie *et al.*, 2001; Poulsen *et al.*, 2010; Li *et al.*, 2018). Ademais, compostos produzidos por uma plethora de microrganismos simbióticos, como as Actinobactérias presentes no tegumento das formigas (encontradas em várias espécies da subtribo Attina), podem atuar contra o crescimento de patógenos especializados (Sen *et al.*, 2009). Em relação aos mecanismos físicos, são conhecidos dois comportamentos defensivos: (a) lambadura do fungo (*grooming*), que remove esporos da superfície do jardim do fungo e os armazena na cavidade oral das formigas, onde também estão presentes actinobactérias (Currie e Stuart, 2001; Little *et al.*, 2006), e (b) capina de fungos (*weeding*), que consiste na remoção de pedaços do jardim que estão infectados (Currie, Stuart, 2001; Little *et al.*, 2006; Nilsson-Moller *et al.*, 2018). Não obstante, essas estratégias usualmente ocorrem após a entrada do patógeno dentro da colônia; e quanto à etapa anterior, de aquisição?

Para reduzir a contaminação por patógenos, formigas cortadeiras (e notavelmente semelhante a outros insetos sociais) usam comportamentos profiláticos para higienizar companheiras de ninho (Cremer *et al.*, 2007; Morelos-Juárez *et al.*, 2010; Reber *et al.*, 2011; de Roode, Lefèvre, 2012; Meunier, 2015). Comportamentos de limpeza de operárias (*allo* e *selfgrooming*) podem ser mais frequentes para aquelas que retornam do forrageamento (Morelos-Juárez *et al.*, 2010; Richard, Errard, 2009). Segundo Richard e Errard (2009), forrageadoras de *Acromyrmex subterraneus* gastam mais tempo fazendo autolimpeza do que operárias não-forrageadoras, provavelmente devido ao seu contato recorrente com contaminantes externos. Além disso, evitar especialmente

ambientes ou operárias contaminadas podem reduzir o risco de infecções (Tranter *et al.*, 2015).

Caso nenhuma forrageadora ou visitante externo trouxer contaminação, a outra porta de acesso à colônia será pela matéria vegetal. Flores e folhas frescas cortadas e coletadas por formigas podem ser potenciais centros de transmissão horizontal de patógenos (Schmid-Hempel, 1998; Griffiths e Hughes, 2010; Van Bael *et al.*, 2009, 2012; Rocha *et al.*, 2017). Devido à dependência do jardim de fungo por tais recursos, há um alto risco de microrganismos antagonistas serem carregados para dentro da colônia junto com esse material (Rocha *et al.*, 2014). Como contra-medida, há um conjunto de comportamentos de preparação e higiene do substrato antes da sua incorporação no jardim de fungo (Quinlan e Cherrett, 1977; Mangone, Currie, 2007). Quinlan e Cherrett (1977) notaram que *Acromyrmex octospinosus* processam o substrato vegetal pela remoção de contaminantes fúngicos da superfície. Além disso, durante o transporte de folhas por uma forrageadora, uma formiga da casta 'mínima' é encontrada sendo carregada junto com o substrato (*caroneira*), possivelmente (a) removendo contaminantes fúngicos da folha antes da sua entrada no formigueiro, e (b) protegendo a forrageadora de forídeos parasitóides (Vieira-Neto *et al.*, 2006; Kitayama *et al.*, 2012).

Vários estudos exploraram as interações entre formigas cortadeiras e fungos endofíticos, comumente encontrados dentro do material foliar que elas cortam. Van Bael e colaboradores (2009) observaram que tais formigas (a) gastam mais tempo processando folhas com altas cargas de endófitos do que aquelas com baixa abundância, e (b) que tal processamento reduz a carga de endófitos antes da inoculação do substrato no jardim do fungo. Além desse controle higiênico, formigas cortadeiras preferem coletar material

foliar contendo baixa abundância de endofíticos, o que pode influenciar em suas preferências de forrageamento (Cobletz; Bael, 2013). Por exemplo, operárias evitam substratos vegetais enriquecidos com *Trichoderma*, um fungo endofítico com potencial antagonista para o fungo mutualista (Rocha *et al.*, 2014, 2017). Desta forma, fica claro a capacidade da imunologia social dessas formigas em policiar os caminhos de entrada e aquisição de microrganismos estranhos, independentemente dos contextos onde essas ameaças são encontradas (i.e., ambiente, alimento, prole ou jardim de fungo, etc.).

3. A memória composta por indivíduos

Assim como o sistema imune composto por células, a *imunologia composta por formigas* lidará contra ameaças que podem colocar em risco a individualidade do superorganismo. Em formigas cortadeiras essa ameaça é duplicada, uma vez que as proteções pela imunologia social implicam duas dimensões: as das formigas e a de seu fungo mutualista. Na seção anterior foi explanado brevemente algumas das estratégias e profilaxias conhecidas nessas formigas. Portanto, finalmente podemos questionar: como essas defesas se integram a uma memória? E como a mesma sugere uma consciência?

3.1 O que é memória imunológica?

Considerada um fenômeno exclusivo do sistema imune adaptativo de vertebrados, a memória imunológica é agora identificada em uma ampla gama de táxons (i.e., microrganismos, plantas, invertebrados e vertebrados) (Pradeu, 2010; Netea *et al.*, 2009). Também é visualizada no sistema imune inato, responsável por respostas rápidas e inespecíficas a antígenos e patógenos, não necessitando de contato prévio (Netea *et*

al., 2009; Pradeu e Pasquier, 2018). Nessa ampla perspectiva, a memória imunológica parece ocorrer por diferentes mecanismos, sejam pelas células e moléculas documentadas na imunidade inata ou pelo sistema CRISPR-Cas (i.e., *priming*; Datsenko *et al.*, 2012).

Os aspectos da memória podem ser entendidos como um fenômeno gradual e multidimensional, variando ao longo de sistemas imunológicos em vários táxons (Pradeu e Pasquier, 2018). As cinco dimensões-chave que compõem uma memória imune são: (i) força, que resulta no aumento da resposta imune em contatos secundários contra o mesmo antígeno/patógeno; (ii) duração, que resulta o tempo máximo que o sistema imune é capaz de elicitar essa resposta aumentada no organismo; (iii) velocidade, compondo o tempo de resposta em encontros secundários; (iv) especificidade, que dita se a resposta intensificada é específica a um determinado antígeno/patógeno; e (v) extinção, ao qual a resposta aumentada deve se dissipar após a eliminação do patógeno e não persistir no organismo. Cada uma dessas dimensões possui gradientes, do mais fraco ao mais forte, do menor ao mais durável, dependendo do sistema imune de cada táxon (Pradeu e Pasquier, 2018). Portanto, é entendível que o fenômeno da memória imune não é estática e semelhante em todos organismos, mas sim variável e composta por diferentes formas.

Sabendo dessa multidimensionalidade e de que colônias de formigas cortadeiras possuem um sistema imune coletivo, podemos hipotetizar que ele carrega uma memória imune? Sim, podemos. E não suporto essa resposta apenas no viés apresentado por Pradeu e Pasquier (2018) para definir memória imune, mas também pelo contexto ecológico que essas formigas e outros insetos sociais estão inseridos. Vamos discutir eles a seguir.

3.2 Indícios de memória imune coletiva

Por que podemos hipotetizar traços de memória imunológica na imunologia social? Estudos ecológicos e evolutivos sugerem que a evolução da memória imune está intrinsecamente relacionada a ambientes ricos em patógenos (Schmid-Hempel, 2003; Graw *et al.*, 2010). A frequência de reinfecção ao longo da vida do hospedeiro ou a diversidade de patógenos encontrados no ambiente, são determinantes evolutivos para otimizar a retenção de informações pelo sistema imune. Como discorrido acima, as formigas cortadeiras são passíveis de entrar em contato com uma alta gama de fungos ao longo da história de vida da colônia. Esse potencial de infecção é alto se considerarmos a coleta de substrato vegetal e a inúmera quantidade de formigas forrageadoras entrando e saindo do ninho. Com base em tais aspectos ecológicos, pode-se esperar que traços de memória imune tenham evoluído para garantir respostas plásticas contra patógenos.

Apresentado o ponto de vista teórico, podemos seguir para o empírico. Existem algumas evidências em insetos sociais que sugerem o fenômeno de memória imune coletiva. Deixo claro que irei apontar exemplos inerentes as formigas cortadeiras e outros que não são provenientes de insetos que cultivam fungos. O primeiro caso é reportado na formiga cortadeira *Acromyrmex echinattior*, ao qual operárias expostas previamente ao fungo entomopatogênico *Metarhizium anisopliae* aumentam o comportamento higiênico em si mesmas e em operárias irmãs quando apresentadas ao mesmo patógeno dias depois (Walker; Hughes, 2009). Algo semelhante foi visto uma década depois por Goes e colaboradores (2022); colônias de *Atta sexdens* intensificaram suas respostas higiênicas contra cinco fungos após serem expostas três vezes a cada um deles, com intervalos de 14 dias entre exposições. Ambos

estudos nos revelam algumas das dimensões que configuram memória imune pela imunologia social (Pradeu; Pasquier, 2018)¹:

- Força: em ambos estudos, de forma coletiva ou individual, as formigas aumentaram o número de comportamentos higiênicos contra o mesmo fungo patógeno previamente exposto. Interessantemente, esse fator foi evidente em dois gêneros diferentes e contra diferentes fungos ascomicetos.
- Extinção: após a limpeza realizada pelas formigas, a resposta imune social declinou (i.e., foi extinta). Ela só foi reativada na exposição futura ao mesmo patógeno.

À parte das formigas que cultivam fungo, temos exemplos na Europa que também sugerem traços de memória imune em formigas. *Formica selysi* quando apresentada a consecutivas exposições aumenta o comportamento de lambedura (Reber *et al.*, 2011). Já a formiga clonal *Platythyrea punctata* tende a aumentar os cuidados para com a sua prole quando expostas previamente a um fungo entomopatógeno, além de passarem mais tempo realizando tal atividade após experiência prévia (Westhus *et al.*, 2014). Por fim, Konrad e colaboradores (2018) notaram que *Lasius neglectus* tende a evitar contato com operárias contaminadas em situações de superinfecções; nessa situação, a colônia sofreu exposição a um patógeno A e dias depois foi infectada novamente, mas com um patógeno B. Nesses exemplos, notamos as dimensões de força (todos os estudos), sugestão de especificidade (em Konrad *et al.*, 2018), extinção, e sugestão de velocidade (em Westhus *et al.*, 2014).

¹ As dimensões de memória imonológica propostas por Pradeu e Pasquier (2018) que não foram citadas, sugerem que ainda faltam estudos que as tenham checado empiricamente em formigas cortadeiras.

Note que, em todos os indícios empíricos citados acima, a imunologia social é sugestiva de *plasticidade*. Em outras palavras, a dinâmica com que a resposta imune da colônia ocorre a um respectivo patógeno pode ser modificada de acordo com experiências prévias. É similar ou análogo ao sistema imune fisiológico, em que a vacina leva ao desenvolvimento de respostas intensas, específicas e rápidas contra o corpo imunizado (memória imunológica ou conhecida também por *imunologia adaptativa*); mas na imunologia social isso é realizado pelas formigas. Portanto, podemos hipotetizar que há uma *correlação* entre a experiência das formigas frente ao patógeno e a resposta higiênica da colônia como um todo. Como resultado, temos a sugestão de memória imune coletiva.

3.3 Como essa memória imune seria suportada?

Apresentados os exemplos empíricos que sugerem traços de memória imune coletiva em formigas cortadeiras, preciso hipotetizar como ela seria formada e mantida considerando premissas biológicas já estabelecidas. Assim, ressalto que as possibilidades discutidas a seguir são de viés *hipotético*, necessitando de abordagens teórico-empírico para validá-las ou não sob a óptica da comunidade científica. Com isso claro, vou basear este tópico sob a seguinte pergunta: *como a memória imune coletiva ocorreria?* Irei citar outros fenômenos já reportados em formigas cultivadoras de fungos que podem fomentar a capacidade da colônia em reter informações sobre infecções prévias.

Iniciaremos com a formação desta suposta ‘memória coletiva’. O primeiro passo para uma formiga aprender (ou *se lembrar*) sobre uma experiência com um fungo maléfico pode depender da sua interação com o jardim de fungo. Essa circunstância seria análoga a do contato entre um

antígeno e células imunológicas. Por meio de uma interação negativa, as formigas poderiam receber sinais dessa ameaça potencial e aonde ela está ocorrendo em sua fungicultura. Esse tipo de comunicação química ocorre por *semioquímicos*. Operárias podem detectar tais produtos químicos liberados pelo jardim de fungo, talvez capazes de informar sobre seu estado de saúde (Green; Kooij, 2018). Além disso, receptores químicos em suas antenas podem detectar compostos orgânicos voláteis liberados por micróbios, principalmente fungos (Yanagawa *et al.*, 2009; Morath *et al.*, 2012). Até agora, como as formigas percebem a saúde de seus jardins de fungos ainda é uma questão sem resposta clara. Uma sugestão pode envolver a transmissão de informações de áreas infectadas da fungicultura, por sinais químicos, até as formigas (Howard *et al.*, 1996; Green; Kooij, 2018; Goes *et al.*, 2020). North e colaboradores (1999) argumentaram duas possíveis formas de comunicação: (i) um sinal volátil liberado pelo jardim de fungo, que se difundiria entre a colônia, ou (ii) um semioquímico liberado em partes danificadas do jardim. Sua configuração experimental mostrou que a primeira hipótese era improvável de acontecer, enquanto a segunda mais plausível de ocorrência. Até o momento, o sinal químico ou a hipótese de informação transmitida pelo próprio jardim de fungo carece de mais evidências experimentais. Embora a comunicação fosse relevante para as formigas captarem informações sobre a situação, eles precisariam vincular com o ‘nível de dano’ ou ‘ameaça’ que o fungo patógeno causou. Durante esse contato, informações sobre o patógeno (quantidade, virulência, pistas químicas; Cremer *et al.*, 2007) e os danos causados por ele seriam adquiridos pelas formigas experientes, unindo ambas em feedback negativo por *aprendizagem associativa* (Goes *et al.*, 2020).

Esse aprendizado por interações sociais mostra notavelmente como experiências podem influenciar no comportamento, pela mera

associação a estímulos ambientais (Leadbeater, Chittka, 2007). No entanto, qualquer aprendizagem associativa exige o armazenamento de informações adquiridas em situações anteriores e sua vinculação aos respectivos estímulos (Josens *et al.*, 2009). As abelhas podem aprender sobre quais espécies de flores são mais gratificantes para forragear, utilizando pistas visuais (cores) em associação com a quantidade de recompensa recebida (Menzel, 1985, 2001; Kevan *et al.*, 1996; Chittka, 1997). No contexto das formigas cortadeiras, é sugerível que uma ‘operária de memória’² utilize a informação armazenada e a utilize para tomada de decisões em uma futura situação semelhante. Para que ocorra o processo de reconhecimento, é necessária a aquisição e retenção de informações pelo indivíduo (Beecher, 1990; Bradbury, Vehrencamp, 1998). Se isso não ocorrer, então a hipótese da memória imune coletiva não será justificada pela união de experiências individuais. Até o momento, algumas pesquisas evidenciaram que formigas cortadeiras são capazes de aprender com feedbacks negativos. Abaixo irei comentar em detalhes esses exemplos.

Ridley e colaboradores (1996) demonstraram que *Atta* e *Acromyrmex* aprendem a rejeitar substratos vegetais que contenham compostos nocivos à sua fungicultura. Aparentemente, as formigas podem perceber mudanças químicas no cultivo de fungos causadas por toxinas de plantas (Knapp *et al.*, 1990; Herz *et al.*, 2008) e param de colhê-las em seguida. Para checar essa capacidade associativa nessas formigas, experimentos em colônias foram feitos com o fungicida cicloheximida, que representaria um composto maléfico ao jardim de fungo (Ridley *et al.*, 1996; North *et al.*, 1999). Após o primeiro oferecimento de iscas com odor de laranja contendo cicloheximida,

² Aquele indivíduo que obteve uma experiência prévia contra um patógeno em potencial.

operárias deixaram de coletar essas iscas após algum tempo. Em seguida, para certificarem que as formigas associaram o dano da cicloheximida com o odor, os autores ofereceram iscas contendo apenas o cheiro de laranja. Surpreendentemente, elas evitaram carregá-los para o jardim de fungo, mesmo sem conter a toxina. Os autores levantam a hipótese de que as formigas associaram o dano causado pela cicloheximida com o odor de laranja (ver mais em North *et al.*, 1997; 1999). Contribuindo, Herz e colegas (2008) aplicaram cicloheximida (com concentrações não detectáveis pelas formigas) em plantas inofensivas ao seu fungo e acompanharam a resposta da colônia. Após dez horas, operárias ingênuas e experientes pararam de carregar as folhas. Em ofertas subsequentes de folhas não tratadas, os formigueiros evitaram colher essa espécie de planta (Herz *et al.*, 2008). Novamente, isso sugere a capacidade do indivíduo em associar o dano em seu cultivar a um determinado espécime, atribuindo a ele um dano potencial. No entanto, a ocorrência dessa plasticidade no comportamento colonial precisa ser testada para com esporos de fungos ou bactérias. Assim, poderemos ter mais certeza se o aprendizado associativo pode explicar o aumento de respostas higiênicas em infecções secundárias ao mesmo patógeno.

A aquisição de informação por indivíduos será relevante para desencadear ações conjuntas rápidas e intensas em infecções futuras. Considere isso como um armazenamento do histórico de infecções desta colônia. Prevemos que uma operária de memória poderia utilizar a infecção prévia para detectar, reconhecer e incitar alertas para o restante da colônia. Como em outros insetos sociais (Cremer *et al.*, 2007), as formigas podem usar feromônios para recrutar companheiras de ninho na aplicação das respostas higiênicas conhecidas (Hölldobler; Wilson, 1990; Gerstner *et al.*, 2011). Assim, tanto operárias experientes

quanto inexperientes seriam recrutadas, aumentando, portanto, a resposta defensiva coletiva em infecções secundárias ao mesmo patógeno. Note que se esse for o caminho mais parsimonioso para explicar a plasticidade na imunologia social, então essa memória coletiva deverá ser *cíclica* durante a história de vida da colônia, pois:

- O tempo com que o aprendizado associativo é capaz de durar/ser retido no cérebro dessas formigas, sob a ausência de reforço ao mesmo estímulo após um tempo. Segundo estudos neurobiológicos realizados por Falibene e colaboradores (2015), informações entre odor da planta e a sua viabilidade para o fungo mutualista duram em torno de 14 dias, enfraquecendo a associação se o estímulo não for ofertado novamente.
- O período de vida de uma operária. Isso poderá variar entre espécies. Lembre-se que uma das principais características de insetos sociais é a sobreposição de gerações. Se o patógeno/estímulo não continuar oferecendo uma ameaça para a colônia por um breve período de tempo, as operárias de memória já terão sido substituídas (ou terem perdido a associação em seu cérebro) quando o patógeno reentrar na colônia.

Considerações finais

É interessante pensarmos em como a memória imune social sustenta a consciência externalizada. A hipótese citada acima é baseada na conjunção de várias experiências individuais frente a uma ameaça, resultando na resposta uníssona de toda a colônia. Então, é uma informação que pode ser acessada pelo todo, afetando assim o

comportamento geral do superorganismo. Como já comentado pela Dra. Gordon (2010), muito do que vemos do comportamento de uma colônia está diretamente relacionada com as experiências de todas as operárias que a compõem. Assim, não podemos esperar o oposto quando analisamos o contexto da memória imune exposto aqui. O que nos falta investigar mais no campo empírico é (i) o envolvimento do aprendizado associativo nessa formação da experiência para com patógenos, (ii) a integração dessas formigas para com a resposta plástica evidenciada em estudos prévios (Walker e Hughes, 2009; Reber et al., 2010; Westhus et al., 2014; Konrad et al., 2018; Goes et al., 2022), e (iii) quanto tempo essa informação é armazenada a nível de colônia. Já ressalto que os traços de memória coletiva poderão sofrer variações entre espécies de formigas cultivadoras de fungos e até mesmo entre insetos sociais. Baseio essa fala no que já visualizamos da evolução do sistema imunológico em diversos táxons (i.e., a multidimensionalidade da memória imune; Pradeu, Pasquier, 2018), e na ecologia dos diferentes insetos sociais ao longo de sua evolução. Não vamos esperar que esse fenômeno possa ser generalizado em qualquer formiga. Mesmo assim, não acho que a ausência dessa memória imune social em algumas espécies invalide a tese da mente estendida. Do contrário, há outros campos da biologia (sem ser a da imunologia) nesses insetos que possam justificar a mesma; e acredito que outros colegas já fizeram isso com exímio em outros capítulos deste livro. Finalizo, portanto, com as principais condições da memória imune coletiva hipotetizadas aqui.

- A capacidade de detectar e reconhecer prováveis invasores nocivos pode depender de interações negativas entre o jardim de fungo e a sinalização deste para as formigas.

- Em hipótese, a 'memória imunológica' da colônia será retida pelo número de interações e experiências adquiridas pelas formigas durante seu tempo de vida, e (ii) ao tempo hábil que essa associação possa ser retida pelo seu sistema nervoso sem reforço. Trata-se, portanto, de uma memória reciclável, sempre necessitando de novas atualizações ao longo da história de vida de uma colônia.
- Cada colônia de formiga cortadeira teria a sua própria memória imunológica, de acordo com a diversidade de microrganismos que a mesma encontrará em seu respectivo meio ambiente. Assim, colônias distintas podem ter informações distintas armazenadas. Também sugiro que essas respostas possam variar entre espécies de insetos sociais.

Referências

- AENEN, D. K; EGGLETON, P., ROULAND-LEFÈVRE C., GULDBERG-FRØSLEV T., ROSENDAHL S.; BOOMSMA J. J. The evolution of fungus-growing termites and their mutualistic fungal symbionts. *PNAS*, v. 99, p. 14887-14892, 2002. DOI: 10.1073/pnas.222313099.
- BEECHER, M. D. The evolution of parental-offspring recognition in swallows. In: *Contemporary Issues in Comparative Psychology*. Eds D. Dewsbury (Sunderland, MA: Sinauer Associates), p. 360-380, 1990.
- BOT, A.N.M; Ortius-Lechner, D.; FINSTER, K.; MAILE, R.; BOOMSMA, J. J. Variable sensitivity of fungi and bacteria to compounds produced by the metapleural glands of leaf-cutting ants. *Insectes Sociaux*, v. 49, p. 363-370, 2002. DOI: 10.1007/PL00012660.
- BOOMSMA, J. J.; GAWNE, R. Superorganismality and caste differentiation as points of no return: how the major evolutionary transitions were lost in translation. *Biological Reviews*, v. 93, p. 28-54, 2018.

- BRADBURY J. W.; VEHRENCAMP, S. L. *Principles of animal communication*. Sunderland, MA: Sinauer Associates, p. 784-785, 1998.
- BUSS, L. *The evolution of individuality*. Princeton: Princeton University Press, 1987.
- COBLETZ K. E.; BAEL, S.A.V. Field colonies of leaf-cutting ants select plant materials containing low abundances of endophytic fungi. *Ecosphere* 4, art 66, 2013. DOI: 10.1890/ES13-00012.1
- CHITTKA, L. Bee color vision is optimal for coding flower color, but flower colors are not optimal for being coded - Why? *Israel. J. Plant Sci.*, v. 45, p. 115-127, 1997.
- CREMER, S.; ARMITAGE, S.A.O.; SCHMID-HEMPEL, P. Social immunity. *Current Biology*, v. 17, p. 693-702, 2007.
- CURRIE, C. R. A community of ants, fungi, and bacteria: a multilateral approach to studying symbiosis. *Annu. Rev. Microbiol.*, v. 55, p. 357-380, 2001. DOI: 10.1146/annurev.micro.55.1.357.
- CREMER S.; SIXT M. Analogies in the evolution of individual and social immunity. *Philos Trans Roy Soc B*, v. 364, p. 129-142, 2009.
- CURRIE, C. R., Mueller, U. G.; Malloch, D. The agricultural pathology of ant fungus gardens. *PNAS.*, v. 96, p. 7998-8002, 1999. DOI: 10.1073/pnas.96.14.7998.
- CURRIE C. R., STUART, A.E. Weeding and grooming of pathogens in agriculture by ants. *Proc. R. Soc. B*. 268, p. 1033-1039, 2001. DOI: 10.1098/rspb.2001.1605.
- DATSENKO, K. A.; POUGACH, K.; TIKHONOV, A.; WANNER, B. L.; SEVERINOV, K.; SEMENOVA, E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.*, v. 3, p. 9-45, 2012.
- FALIBENE A.; ROCES, F.; ROSSLER, W. Long-term avoidance memory formation is associated with a transient increase in mushroom body synaptic complexes in leaf-cutting ants. *Front. Behav. Neurosci.*, p. 9-84, 2015.
- FERNÁNDEZ-MARÍN, H.; ZIMMERMAN, J. K.; REHNER, S. A.; WCISLO, W. T. Active use of the metapleural glands by ants in controlling fungal infection. *Proc. R. Soc. B*. 273, p. 1689-1695, 2006. DOI: 10.1098/rspb.2006.3492.

- FERNÁNDEZ-MARÍN H, *et al.* Functional role of phenylacetic acid from metapleural gland secretions in controlling fungal pathogens in evolutionarily derived leaf-cutting ants. *Proc. R. Soc. B.* 282, p. 1807, 2015. DOI: 10.1098/rspb.2015.0212.
- GARRETT W. S.; GORDON J. I.; GLIMCHER, L. H. Homeostasis and inflammation in the intestine. *Cell*, v. 140, n. 6, p. 859–870, 2010.
- GERSTNER, A. T.; POULSEN; M. CURRIE, C. R. Recruitment of minor workers for defense against a specialized parasite of *Atta* leaf-cutting ant fungus gardens. *Ethol. Ecol. Evol.* v. 23, p. 61–75, 2011.
- GOES A. C.; BARCOTO M.O.; KOOIJ P. W.; BUENO O. C.; RODRIGUES, A. How do leaf-cutting ants recognize antagonistic microbes in their fungal crop? *Frontiers in Ecology and Evolution*, p. 8-95, 2020.
- GOES, A. C.; KOOIJ, P. W.; CULOT, L.; BUENO, O. C.; RODRIGUES, A. Distinct and enhanced hygienic responses of a leaf-cutting ant toward repeated fungi exposures. *Ecol. Evol.* v. 12, e9112, 2022.
- GOULD, S. J.; LLOYD, E. Individuality and adaptation across levels of selection: How shall we name and generalize the unit of Darwinism? *Proceedings of the National Academy of Sciences USA* v. 96, n. 21, p. 11904–11909, 1999.
- GORDON, D. M. *Ant Encounters: Interaction Networks and Colony Behavior*. Princeton University, 2010.
- GRAYSTOCK, P.; HUGHES, W. O. H. Disease resistance in a weaver ant, *Polyrhachis dives*, and the role of antibiotic-producing glands. *Behav. Ecol. and Socio.* v. 65, p. 2319–2327, 2011. DOI: 10.1007/s00265-011-1242-y.
- GRAW, F., MAGNUS, C., REGOES, R. R. Theoretical analysis of the evolution of immune memory. *BMC Evol. Biol.*, v. 10, n. 380, 2010.
- GREEN, P. W. C.; Kooij, P. W. (2018). The role of chemical signaling in maintenance of the fungus garden by leaf-cutting ants. *Chemoec*, 2018. DOI: 10.1007/s00049-018-0260-x.
- GRIFFITHS, H. M.; HUGHES, W. O. H. Hitchhiking and the removal of microbial contaminants by the leaf-cutting ant *Atta colombica*. *Ecol Entomol.* v. 35, p. 529–537, 2010. DOI: 10.1111/j.1365-2311.2010.01212.x.

- HERZ, H.; HÖLLDOBLER, B.; ROCES, F. Delayed rejection in a leaf-cutting ant after foraging on plants unsuitable for the symbiotic fungus. *Behav. Ecol.* v. 19, p. 575–582, 2008.
- HÖLLDOBLER, B. On the metapleural gland of ants. *Psyche.* v. 91, p. 91201–91224, 1984. DOI: 10.1155/1984/70141.
- HÖLLDOBLER, B.; WILSON, E. O. *The Ants*. Cambridge, MA, Harvard University Press, 1990.
- HOWARD, J. J. L.; HENNEMAN, C. G.; CRONIN, J. A. F.; HORMIGA, G. Conditioning of scouts and recruits during foraging by leaf-cutting ant, *Atta colombica*. *Anim. Behav.* v. 52, p. 299–306, 1996.
- JOSENS, R.; ESCHBACH, C.; GIURFA, M. Differential conditioning and long-term olfactory memory in individual *Camponotus fellah* ants. *Journ. of Exper. Biol.* v. 212, p. 1904–1911, 2009.
- KEVAN, P.; GIURFA, M.; CHITTKA, L. Why are there so many and so few white flowers? *Trends Plant. Sci.* 1, p. 280–284, 1996.
- KITAYAMA, K.; SOUSA-SOUTO, L.; AQUINO, P. P. U.; TENORIO, L. X. Trichome removal by hitchhikers in two leaf-cutting ant species (Hymenoptera: Formicidae). *Sociobiology*, v. 59, p. 935–943, 2012.
- KONRAD, M. *et al.* Ants avoid superinfections by performing risk-adjusted sanitary care. *Proc. Natl. Acad. Sci. U.S.A.*, v. 115, 2782–2787, 2018
- MARTIN, M. M.; MACCONNELL, J. G.; GALE, G. R. The chemical basis for the attine ant-fungus symbiosis. Absence of antibiotics. *Ann. of the Entomo. Soci. of Amer.* v. 62, p. 386–388, 1969. DOI: doi.org/10.1093/aesa/62.2.386.
- MICHOD, R. E. *Darwinian dynamics: evolutionary transitions in fitness and individuality*, Princeton, Princeton University Press, 1999.
- MUELLER, U. G. Ant versus fungus versus mutualism: ant-cultivar conflict and the deconstruction of the attine ant-fungus symbiosis. *American Naturalist*, v. 160, p. 67–98, 2002. DOI: 10.1086/342084.

- MOLLER, A. Die Piltzgarten einiger sudamerikanischer Ameisen. *Botan. Mittei. Tropen*, v. 6, p. 1-127, 1983.
- PULL, C; MCMAHON, D. *Superorganism immunity*: a major transition in immune system evolution. *Frontiers in Ecology and Evolution*, p. 8-186, 2020.
- LEADBEATER, E.; CHITTKA, L. The dynamics of social learning in an insect model, the bumblebee (*Bombus terrestris*). *Behav. Ecol. Sociobiol.*, v. 11, p. 1789-1796, 2007.
- LI, H.; SOSA-CALVO, J.; HORN, H. A.; PUPO, M. P.; CLARDY, J.; RABELING, C.; SCHULTZ, T. R. Convergent evolution of complex structures for ant – bacterial defensive symbiosis in fungus-farming ants. *PNAS*, p. 1–6, 2018. DOI: 10.1073/pnas.1809332115.
- LITTLE, A. E. F.; MURAKAMI, T.; MUELLER, U. G.; CURRIE, C. R. Defending against parasites: fungus-growing ants combine specialized behaviours and microbial symbionts to protect their fungus gardens. *Biol Lett.*, v. 2, p. 12–16, 2006. doi: 10.1098/rsbl.2005.0371.
- MANGONE D. M.; CURRIE, C. R. Garden substrate preparation behaviours in fungus-growing ants. *Can. Entomol.*, v. 139, p. 841-849, 2007. DOI: 10.4039/n06-105.
- MENZEL, R. Learning in honey bees in an ecological and behavioral context. *Fortschritte der Zoologie*, v. 31, p. 55–74, 1985.
- MENZEL, R. Behavioral and neural mechanisms of learning and memory as determinants of flower constancy. In: *Cognitive Ecology of Pollination: Animal Behaviour and Floral Evolution*, eds L. Chittka and J.D. Thomson Cambridge: Cambridge University Press, p. 21–40, 2001.
- MEUNIER, J. Social immunity and the evolution of group living in insects. *Phil. Trans. R. Soc. B.* 370, 20140102, 2015. DOI: 10.1098/rstb.2014.0102.
- MORELOS-JUÁREZ, C.; WALKER, T.N.; LOPES, J. F. S.; HUGHES, O. H. W. Ant farmers practice proactive personal hygiene to protect their fungus crop. *Curr. Biol.*, v. 20, n. 13, R553-4. p. 553–54, 2010. DOI: 10.1016/j.cub.2010.04.047.
- MORATH, S. U.; HUNG, R.; BENNETT, J. W. Fungal volatile organic compounds: a review with emphasis on their biotechnological potential. *Fungal Biol Rev.* 26, p. 73–83, 2012. DOI: 10.1016/j.fbr.2012.07.001.

- NETEA, M. G.; QUINTIN, J.; VAN DER MEER, J. W. M. Trained immunity: a memory for innate host defense. *Cell host microbe*, v. 9, p. 355–361, 2011.
- NILSSON-MOLLER, S.; POULSEN, M.; INNOCENT, T. M. A visual guide for studying behavioral defenses to pathogen attacks in leaf-cutting ants. *J. Vis. Exp.* v. 140, e58420, 2018. DOI: 10.3791/58420.
- NORTH, R. D.; JACKSON, C. W.; HOWSE, P. E. Evolutionary aspects of ant–fungus interactions in leaf-cutting ants. *Trends Ecol. Evol.*, v. 12, p. 386–389, 1997.
- NORTH, R. D.; JACKSON, C. W.; HOWSE, P. E. Communication between the fungus garden and workers of the leaf-cutting ant, *Atta sexdens rubropilosa*, regarding choice of substrate for the fungus. *Physiol. Entomol.*, v. 24, p. 127–133, 1999.
- OTANI, S., et al. Disease-free monoculture farming by fungus-growing termites. *Scien. Reports*, v. 9, 8819, 2019. DOI: 10.1038/s41598-019-45364-z.
- ORTIUS-LECHNER, D.; MAILE, R.; MORGAN, E. D.; BOOMSMA, J. J. Metaplural gland secretion of the leaf-cutter ant *Acromyrmex octospinosus*: new compounds and their functional significance. *Jour. of Chem. Ecol.* 26, p. 1667–1683, 2000. DOI: 10.1023/A:1005543030518.
- POULSEN, M.; BOT, A. N. M.; NIELSEN, M. G.; BOOMSMA, J. J. Experimental evidence for the costs and hygienic significance of the antibiotic metapleural gland secretion in leaf-cutting ants. *Behav. Ecol. and Sociob.* 52, p. 151–157, 2002. DOI: 10.1007/s00265-002-0489-8.
- POULSEN, M., et al. Variation in *Pseudonocardia* antibiotic defence helps govern parasite-induced morbidity in *Acromyrmex* leaf-cutting ants. *Envi. Micro. Repor*, v. 2, p. 534–540, 2010. DOI: 10.1111/j.1758-2229.2009.00098.x.
- PRADEU, T. *The Limits of the self: Immunology and biological identity*. New York: Oxford University Press, 2012.
- PRADEU, T. A mixed self: the role of symbiosis in development. *Biol. Theory*, v. 6, n. 1, 2011.
- PRADEU T; PASQUIER, L. D. Immunological memory: what's in a name? *Immun. Rev.*, v. 283, p. 7–20, 2018.

- SCHMID-HEMPEL, P. *Parasites in social insects*. Princeton, EUA: Princeton University Press, 1998.
- QUINLAN, R. J.; CHERRETT, J. M. Role of substrate preparation in symbiosis between leaf-cutting ant *Acromyrmex octospinosus* (Reich) and its food fungus. *Ecol. Entomol.*, v. 2, p. 161-170, 1977. DOI: 10.1111/j.1365-2311.1977.tb00877.x.
- REBER, A.; PURCELL, J.; BUECHEL, S. D.; BURI, P.; CHAPUISAT, M. The expression and impact of antifungal grooming in ants. *J Evol Biol.*, v. 24, p. 954–964, 2011. DOI: 10.1111/j.1420-9101.2011.02230.x.
- RICHARD, F.; ERRARD, C. Hygienic behavior, liquid-foraging, and trophallaxis in the leaf-cutting ants, *Acromyrmex subterraneus* and *Acromyrmex octospinosus*. *Journ. of Insec. Sci.*, v. 9, p. 1–9, 2009. DOI: 0.1673/031.009.6301.
- RIDLEY P.; HOWSE P. E.; JACKSON, C. W. Control of the behaviour of leaf-cutting ants by their ‘symbiotic’ fungus. *Experientia*, v. 52, p. 631–635, 1996.
- DE ROODE, J. C.; LEFÈVRE, T. Behavioral immunity in insects. *Insects*, v. 3, p. 789–820, 2012. DOI: 10.3390/insects3030789.
- ROCHA, S. L.; JORGE, V. L.; DELLA LUCIA, T. M. C.; BARRETO, R. W.; EVANS, H. C.; ELLIOT, S. L. Quality control by leaf-cutting ants: evidence from communities of endophytic fungi in foraged and rejected vegetation. *Art. Plant Int.*, v. 8, p. 485–493, 2014. DOI:10.1007/ s11829-014-9329-9.
- ROCHA S. L.; EVANS, H. C.; JORGE, V. L. ; CARDOSO, L. A. O.; *et al.* Recognition of endophytic *Trichoderma* species by leaf-cutting ants and their potential in a Trojan-horse management strategy. *R. Soc. Open Sci.*, v . 4, 160628, 2017. DOI: 10.1098/rsos.160628.
- RODRIGUES, A.; BACCI Jr., M.; MUELLER, U. G.; ORTIZ, A.; PAGNOCCA, F. C. Microfungal “weeds” in the leafcutter ant symbiosis. *Micro. Ecol.*, v. 56, p. 604–614, 2008. DOI: 10.1007/s00248-008-9380-0.
- RODRIGUES, A.; MUELLER, U. G.; ISHAK, H. D.; BACCI Jr., M.; PAGNOCCA, F. C. Ecology of microfungal communities in gardens of fungus-growing ants (*Hymenoptera*:

- Formicidae*): a year-long survey of three species of attine ants in Central Texas. *FEMS Microb. Ecol.*, v. 78, p. 244–255, 2011. DOI: 10.1111/j.1574-6941.2011.01152.x.
- SCHMID-HEMPEL, P. Variation in immune defence as a question of evolutionary ecology. *Proc. Biol. Sci. B.*, v. 270, p. 357–366, 2003.
- SEN, R.; ISHAKA, H.; ESTRADA, D.; DOWD, S.; HONGA, E.; MUELLER, U. Generalized antifungal activity and 454-screening of *Pseudonocardia* and *Amycolatopsis* bacteria in nests of fungus-growing ants. *PNAS.*, v. 106, 17805, 2009. DOI: 10.1073/pnas.0904827106.
- SILVA, A.; BACCI Jr., M.; PAGNOCCA, F. C.; BUENO, O. C.; HEBLING, M. J. A. Starch metabolism in *Leucoagaricus gongylophorus*, the symbiotic fungus of leaf-cutting ants. *Microbi. Res.*, v. 161, p. 299–303, 2006. DOI: 10.1016/j.micres.2005.11.001.
- SIQUEIRA, G. C. de; BACCI Jr., M.; PAGNOCCA, F. C.; BUENO, O. C.; HEBLING, M. J. A. Metabolism of plant polysaccharides by *Leucoagaricus gongylophorus*, the symbiotic fungus of the leaf-cutting ant *Atta sexdens*. *App. and Enviro. Micro.*, v. 64, p. 4820–4822, 1998. DOI: 10.1016/j.micres.2005.11.001.
- STRASSMANN, J. E.; QUELLER, D. C. Genetic conflicts and intercellular heterogeneity. *Journal of Evolutionary Biology*, v. 17, p. 1189–1191, 2004.
- TRANTER, C.; LEFÈVRE, L.; EVISON, SEF.; HUGHES, W. O. H. Threat detection: contextual recognition and response to parasites by ants. *Behav. Ecol.* 26, 396–405, 2015. DOI: 10.1093/beheco/aru203.
- TRIVIÑO, V.; SUÁREZ, J. Holobionts: Ecological communities, hybrids, or biological individuals? A metaphysical perspective on multispecies systems. *Studies in History and Philosophy of Biol. & Biomed. Sci.*, v. 84, n. 101323, 2020.
- VAN BAEL, S. A.; FERNÁNDEZ-MARÍN, H.; VALENCIA, M. C.; ROJAS, E. I.; WCISLO, W. T.; HERRE, E. A. Two fungal symbioses collide: endophytic fungi are not welcome in leaf-cutting ant gardens. *Proc. Biol. Sci.*, v. 276, p. 2419–2426, 2009. DOI: 10.1098/rspb.2009.0196.

- VAN BAEEL, S. A.; SEID, M.; WCISLO, W. Endophytic fungi increase the processing rate of leaves by leaf-cutting ants (*Atta*). *Ecol. Entomol.*, v. 37, p. 318-321, 2012. DOI: 10.1111/j.1365-2311.2012.01364.x.
- VIEIRA-NETO, E. H. M.; MUNDIN, F. M.; VASCONCELOS, H. L. Hitchhiking behaviour in leaf-cutter ants: an experimental evaluation of three hypotheses. *Insectes Sociaux*, v. 53, p. 326-332, 2006. DOI: 10.1007/s00040-006-0876-7.
- YANAGAWA, A.; YOKOHARI, F.; SHIMIZU, S. The role of antennae in removing entomopathogenic fungi from cuticle of the termite, *Coptotermes formosanus*. *J. Insect Sci.* v. 9, n. 6, 2009. DOI: 10.1673/031.009.0601.
- WALKER N. T.; HUGHES WHO. Adaptative social immunity in leaf-cutting ants. *Biol. Lett.*, v. 5, p. 446-448, 2009.
- WESTHUS, C.; UGELVIG, L. V.; TOURDOT, E.; HEINZE J.; DOUMS C, CREMER, S. Increased grooming after repeated brood care provides sanitary benefits in a clonal ant. *Behav. Ecol. Sociobiol.*, v. 68, p. 1701-1710, 2014.
- WEBER, N. A. The fungus-culturing behavior of ants. *Am. Zool.*, v. 12, p. 577-587, 1972.
- WHEELER, W. M. The ant-colony as an organism. *J. Morph.*, v. 22, p. 307-325, 1911.

5

MEMÓRIA COLETIVA E CONSCIÊNCIA COLETIVA SÃO FENÔMENOS CIENTÍFICOS?

José Carlos Camillo

Os estudos científicos sobre a memória começaram na virada do século XIX para o século XX. Enquanto as humanidades desde então se focaram em estudar a memória coletiva—filósofos inclusos nesses estudos—apenas a partir da década de 1980 que a memória coletiva tomou o interesse de cientistas cognitivos. Não é sem críticas, no entanto, que tais estudos foram realizados.

Neste capítulo, tenho um objetivo duplo. Primeiro, defendo que a memória coletiva é um fenômeno científico e, portanto, tem validade enquanto objeto de estudo da ciência. Adoto uma abordagem mais neutra quanto à realidade dos fenômenos científicos neste capítulo. Ou seja, apesar de sua validade enquanto objeto de estudo, não defendo que ele seja um fenômeno objetivo, pois isso exigiria uma defesa da postura realista quanto à prática científica, o que foge do escopo deste capítulo. Consequentemente, ao defender a validade da memória coletiva enquanto fenômeno científico, realistas podem concluir que tal fenômeno existe de fato, independente da teorização científica, e antirrealistas podem concluir que sua postulação é importante diante das práticas científicas. De qualquer modo, tal fenômeno não deveria ser descartado das investigações científicas. Uma importante clarificação é que me focarei em memória de grupos pequenos e não de grupos grandes, como países ou etnias inteiras¹.

¹ A razão de meu foco é simplesmente de familiaridade com a literatura. Alguns estudos já demonstraram que elas se comportam de forma diferente (Michaelian; Sutton, 2019; Arango-Muñoz; Michaelian, 2020). Portanto, abordar as duas em um único capítulo tornaria o capítulo muito longo e

Em segundo lugar, após defender que a memória coletiva é um fenômeno científico, discuto as possibilidades de inferir de modo confiável, a partir disso, um fenômeno de consciência coletiva. Schwitzgebel (2015), por exemplo, defende que a organização e atividade de certos grupos de uma certa forma que favorece a atribuição de capacidades cognitivas a eles implica na atribuição de consciência coletiva. Argumento que tal inferência confiável é improvável e que, mesmo que se aceite a confiabilidade de tal inferência, no estilo de Schwitzgebel, as coisas são muito mais complicadas do que a mera atribuição de memória coletiva a grupos.

Para alcançar meus objetivos, começo apresentando o conceito de fenômeno científico adotado aqui e sua neutralidade ontológica (Seção 1). Depois, argumento que a memória coletiva satisfaz o conceito de fenômeno científico apresentado e, portanto, é um fenômeno válido e digno de ser investigado (Seção 2). Dada essa conclusão, considero o argumento de que capacidades cognitivas atribuídas a grupos implicam na atribuição de consciência coletiva (Seção 3). Apesar de indicar que tal inferência ou implicação é improvável, defendo que, mesmo se for, as coisas são mais complicadas e a inferência deve levar em conta o tipo de memória atribuído à coletividade e a dimensão de consciência que se pode inferir de tal tipo de memória.

1. Sobre fenômenos científicos

Meu objetivo neste capítulo é abordar uma questão científica e não ontológica. Isto é, ao invés de defender que existe uma memória coletiva independente da teorização humana, defenderei que memória coletiva

complexo. Por isso, foquei-me no tipo de memória coletiva com o qual tenho mais familiaridade com a literatura especializada.

é um fenômeno científico, tendo importância para a teorização humana sobre memória (questão científica). Como tal, realistas terão razão para acreditar que tal fenômeno é real (questão ontológica). No entanto, não adentrarei na defesa do realismo, pois tal debate foge ao escopo de minha argumentação. Um exemplo da dissociação entre fenômeno científico e fenômeno real está nos estudos dos elétrons e outras partículas e fenômenos subatômicos. A postulação da existência deles é crucial para a teorização da física de partículas. No entanto, há um intenso debate a respeito de se tal postulação implica na postulação de entidades reais, independentes da teorização, ou apenas entidades teóricas, dependentes da teorização. Apesar de pessoalmente favorecer propostas realistas, minha argumentação ao longo deste capítulo adota uma posição neutra em relação a essa questão.

Na linguagem corriqueira, fenômeno normalmente se refere a eventos. Na prática científica, um fenômeno é o objeto de investigação científica² e se define por um conjunto estável de propriedades. Esse conjunto aparece sob diversas naturezas—evento, processo, estado ou objeto (Bogen & Woodward, 1988, Brechtel, 2007, Craver & Darden, 2013, Colaço, 2018). Essas propriedades, em geral, são causais, pois sua detecção é realizada através de inferências causais a partir de experimentação. Ou seja, propriedades causais são passíveis de serem mensuradas e são relevantes para a individuação de fenômenos por causa de seu poder causal (Harinen, 2018). Esse conjunto estável de propriedades causais tende a coocorrer dadas certas condições. Por isso, uma forma muito comum de individuação e caracterização de um

² A ciência se propõe a estudar vários tópicos (Potochnik, 2017), no entanto a maioria de seus esforços se restringem a investigar para explicar fenômenos (Bogen & Woodward, 1988, Woodward, 2003). Por isso, ao longo do texto considero como fenômeno científico aquilo que é objeto de investigação científica.

fenômeno envolve a descoberta das condições nas quais esse conjunto emerge e se estabiliza (Craver & Darden, 2013, Colaço, 2018). No entanto, a ciência também investiga fenômenos de ocorrência única, como o *Big Bang*. Nesses casos, a descoberta do fenômeno é feita através dos efeitos de suas propriedades causais, como o movimento de expansão dos corpos celestes e a radiação cósmica de fundo (Brechtel, 2007).

No caso de fenômenos recorrentes, cientistas notam uma certa constância ou estabilidade de propriedades. Por exemplo, estudos científicos da memória começaram quando Hermann Ebbinghaus realizou alguns testes consigo mesmo (Baddeley *et al*, 2020). Para evitar a influência de aprendizados prévios, o filósofo e psicólogo alemão criou muitas listas de não-palavras, como *zis*, para decorar. Ele contabilizou o tempo e a acurácia de várias tentativas diárias de decorar listas ao longo de mais ou menos dois anos. Com isso, ele percebeu que existiam formas diferentes de estabilidade da informação sobre listas na mente³.

Por um lado, se fosse apresentada uma lista curta de não-palavras, de menos de cinco itens, e ele fosse testado logo em seguida, ele teria uma acurácia boa e o tempo de retomada dos itens seria rápido. No entanto, essa lista seria drasticamente esquecida em questão de segundos. A mesma coisa não acontecia com listas grandes, de mais de cinco itens. Dessa forma, tinha-se um fenômeno: itens eram

³ Os estudos de Ebbinghaus constituem uma exceção na história da ciência (ver Baddeley *et al*, 2020). Ele mesmo foi o criador, aplicador e objeto de estudo dos testes. Ele criou uma série enorme de testes diferentes, com diferentes quantidades de palavras e diferentes métodos para memorização. A partir disso, ele calculou com precisão não apenas quanto tempo ele levava para esquecer as palavras considerando cada método—ou seja, a velocidade do esquecimento—, mas também a quantidade de palavras que tendiam a permanecer ou se esvaziar. Esses estudos são uma exceção pois a condição de formulador, aplicador e sujeito dos testes tende a criar vieses nos resultados. No entanto, quando aplicados em situações mais controladas—duplo cego, por exemplo—os resultados de Ebbinghaus foram consistentemente replicados. Além disso, como será apresentado, métodos diferentes também conduziram a resultados semelhantes. Isso fortalece a proposta de que Ebbinghaus, de fato, encontrou fenômenos científicos.

rapidamente memorizados, com alta acurácia e eram rapidamente esquecidos. Esse conjunto de propriedades estáveis ocorria quando era apresentada uma lista curta, mas não quando era apresentada uma lista grande, o que implicava que o número de itens decorados era pequeno. Em suma, o fenômeno descoberto tinha as seguintes propriedades: rápida memorização, rápido e abrupto esquecimento, alta acurácia e baixo número de itens. Essas propriedades surgiam quando havia uma apresentação rápida dos itens, seguida de uma retomada (teste) quase imediata, e quando a lista apresentada era pequena. Em outras condições, o fenômeno não ocorria.

Por outro lado, Ebbinghaus notou um fenômeno diferente em seus estudos. Se houvesse repetidas apresentações de uma lista (ou seja, se ele visse a lista várias vezes e se esforçasse várias vezes para decorá-la) e se fosse seguido um padrão de estudo-descanso, listas bem maiores poderiam ser memorizadas. Além disso, o tempo de esquecimento, nesses casos, era muito maior (de dias a semanas), consistentemente relacionado ao número de apresentações/estudos e da prática de intervalo/descanso entre elas. Interessantemente, não parecia haver limites para o número itens de uma lista, nem para o número de listas, já que decorar uma lista não impedia a memorização de outra. Esse segundo fenômeno tem como propriedades: memorização lenta e esquecimento lento e gradual, relacionado ao número de apresentações e aos intervalos, boa acurácia e número indeterminado de itens. As condições para a ocorrência de fenômeno eram múltiplas apresentações da lista com intervalos de descanso entre essas apresentações.

Esses dois fenômenos são alvos de investigações científicas de sucesso e são cruciais para a teorização da memória humana. Eles são fenômenos individuados e caracterizados de formas diferentes devido a suas distintas propriedades estáveis e às distintas condições nas quais

emergem. Atualmente, esses fenômenos são conhecidos como memória de curto e longo prazo;

Tomando o exemplo desses fenômenos, entende-se que um conjunto de propriedades é um fenômeno científico se essas propriedades tendem a coocorrer em condições similares—na prática científica, um fenômeno recorrente deve ser *replicável*. A replicabilidade se dá quando, mantendo as condições típicas do fenômeno inicialmente encontrado, testa-se se os resultados encontrados inicialmente se repetem (Bogen & Woodward, 1988, Colaço, 2018). Ou seja, ao manterem-se as condições relevantes (apresentação rápida dos itens, seguida de uma retomada (teste) quase imediata, e uma lista pequena de itens), encontra-se as propriedades tidas como estáveis que constituem o fenômeno (rápida memorização, rápido e abrupto esquecimento, alta acurácia e baixo número de itens).

De fato, fenômenos que falham na replicabilidade são abandonados enquanto objeto de investigação científica, perdendo o status de objeto de investigação científica (Colaço, 2018, 2020). Por exemplo, fenômenos paranormais não são fenômenos científicos e, portanto, não são alvo de investigação científica (Bogen & Woodward, 1988) porque falham na replicabilidade. Falhas de replicabilidade indicam que não há um conjunto *estável* de propriedades e, portanto, não há um fenômeno válido a ser investigado pela ciência. No entanto, por motivos óbvios, o critério da replicabilidade não se aplica a fenômenos de ocorrência única. Por isso, outro critério é necessário para a validação de um fenômeno enquanto alvo de investigação científica.

Esse segundo critério é a *suscetibilidade à investigação sob diversas metodologias* (Bogen & Woodward, 1988, Brechtel, 2007). Tal critério, assim como o da replicabilidade, advém da necessidade da *estabilidade* das propriedades que compõem o fenômeno. Se há uma estabilidade, tais

propriedades podem ser investigadas de diversas formas. No caso do *Big Bang*, como mencionado, sua investigação é feita de diversas formas, incluindo o uso do efeito Doppler para a observação do movimento das galáxias e a detecção de ondas de rádio da radiação cósmica de fundo. Fenômenos paranormais, em geral, também falham na satisfação desse critério e, portanto, não são alvo de investigação científica. Considerando as memórias de curto prazo e longo prazo, não só testes de memorização com não-palavras ao estilo de Ebbinghaus, mas também testes com palavras, com itens não-verbais (desenhos), testes de rememoração livre ou de reconhecimento e investigação neurobiológica estabeleceram ambos como fenômenos válidos enquanto objeto de investigação científica⁴ (Baddeley *et al*, 2020, Kandel, 2006).

De fato, deve-se ter cautela em considerar que fenômenos científicos são fenômenos reais. Por exemplo, Colaço (2018, 2020) analisa casos em que algo havia sido estabelecido enquanto fenômeno científico a partir dos dois critérios apresentados e, após investigações mais aprofundadas, notou-se que aquele conjunto de propriedades falhava em satisfazer os dois critérios. Por causa disso, deve-se manter cautela em atribuir existência a um fenômeno científico pois, posteriormente, ele pode ser rejeitado ou recharacterizado (o que os cientistas consideravam ser um fenômeno, na verdade, era outro fenômeno). De fato, existem atualmente dados que põem em xeque a caracterização das memórias de curto e longo prazo conforme feita acima (Gershman, 2023).

Além disso, muitas coisas (processos, eventos ou objetos) existentes não se qualificam enquanto fenômeno científico. Bogen e Woodward

⁴ Um adendo muito importante que justifica a neutralidade adotada neste capítulo é que a validação de um conjunto de propriedades como fenômeno científico não implica em sua existência e o reverso também é verdadeiro, a não validação não implica em não existência. Apesar disso, a validação de algo enquanto fenômeno científico indica sua importância para a teorização, especialmente porque fenômenos são usados para validar, explicar ou criticar teorias científicas (Bogen & Woodward, 1988).

(1988) argumentam que a ciência não investiga dados, mas investiga fenômenos que são inferidos ou percebidos a partir desses dados. Por exemplo, entende-se que a água ferve a cem graus Celsius em condições adequadas de temperatura e pressão. O ponto de ebulição da água, é um fenômeno a ser investigado pela ciência. No entanto, se um cientista utilizar muitos termômetros bastante precisos do mesmo tipo, no mesmo ambiente e ao mesmo tempo é provável e esperado que cada um deles apresente leves variações (na casa dos centésimos ou milésimos, por exemplo) e é possível que nenhum deles de fato indique a ebulição da água a exatos 100°C. Essas diversas variações não são objeto de estudo da ciência, pois dependem de inúmeros fatores que fogem do controle dos cientistas e mesmo quando se investigam variações, se selecionam variações devido a fatores específicos e, portanto, generalizáveis. Em outras palavras, ainda assim não se investiga dados, mas fenômenos inferidos ou percebidos a partir desses dados. Isso não quer dizer que os eventos/processos/objetos individuais que resultam nesses dados variáveis não sejam existentes. Ao contrário, ninguém duvidaria da existência deles, a não ser idealistas e céticos mais radicais que Berkeley. Apesar disso, não se validam enquanto fenômeno científico⁵.

Em suma, um fenômeno científico é um conjunto estável de propriedades que emerge sob condições típicas⁶. Os critérios para a validação de um conjunto de propriedades enquanto fenômeno científico são sua replicabilidade e suscetibilidade a diversas

⁵ Como mencionei fenômenos paranormais ao longo dessa seção como exemplo de algo que não se constitui como fenômeno científico, vale mencionar que, também eles podem ser existentes apesar de não serem válidos para a investigação científica. A questão seria como provar ou argumentar em favor de sua existência não utilizando-se da investigação científica.

⁶ Kaiser e Krickel (2016) convincentemente argumentam que as propriedades causais de um fenômeno devem resultar de interações entre entidades físicas (objeto) e devem ser consideradas como resultado da ação desses objetos, quando organizados de determinada forma. No entanto, para não alongar muito o capítulo, deixo de lado essa argumentação, reconhecendo-a nesta nota.

metodologias de investigação científica. Na próxima seção, argumento que a memória coletiva satisfaz esses e outros critérios.

2. Memória coletiva enquanto fenômeno científico

Em geral, argumentos a favor da memória coletiva seguem percursos duvidosos e dependentes de teorias incertas⁷. Por exemplo, Theiner e colegas (2010) argumentam que não há bons argumentos contrários à existência de memória coletiva e que, portanto, tal fenômeno deve ser aceito como real—eles simplesmente mudaram o ônus da prova. Michaelian e Sutton (2019) precisam considerar que memória é uma propriedade disposicional que tende a produzir comportamentos que são reconhecidos socialmente como mnêmicos. Theiner (2013) adota um princípio de paridade funcional para validar seu argumento. Arango-Muñoz e Michaelian (2020) se baseiam em critérios essencialistas de mnemicidade para defender a existência de memória coletiva. Todas essas teorias/metodologias adotadas, no entanto, são bastante incertas, especialmente tendo em vista que o objetivo deles é defender que memória coletiva é um fenômeno real, não meramente um fenômeno científico.

Por ter uma proposta de argumentação mais modesta—defender que memória coletiva é um fenômeno científico—me valerei apenas daquilo que é necessário para tal, ou seja, de critérios que façam sentido dentro da prática científica. No entanto, memória coletiva não se coloca como um fenômeno qualquer. Ela se coloca como um fenômeno coletivo e mnêmico. Dessa forma, além de avaliar se memória coletiva é um fenômeno científico, deve-se avaliar se ela é, de fato, coletiva e

⁷ Não que tais teorias estejam erradas, mas elas estão longe de serem unanimidade. Dessa forma, se alguém discorda da teoria adotada para construir o argumento, automaticamente a conclusão do argumento não convencerá esse alguém.

mnêmica. Para isso, considerarei, além dos critérios para validação de algo enquanto fenômeno científico, dois critérios bem simples:

1. *Critério de fenômeno coletivo*: o fenômeno deve depender de uma coletividade e deve ser distinto de fenômenos de memória individual.
2. *Critério de mnemicidade*: o fenômeno deve se associar a uma mudança de comportamento/performance devido a experiências prévias.

O motivo para adotar o Critério 1 é que a postulação de um fenômeno só faz sentido na prática científica se ele for distinto de fenômenos já estabelecidos, como os fenômenos de memória individual. Essa diferença, por ser um fenômeno coletivo, deve se dar, em parte, em virtude de acontecer em condições coletivas. Entendo condições coletivas aqui como as situações em que mais de um indivíduo interagem entre si. A razão para adotar o Critério 2 é que ele se vale de uma definição abrangente de mnemicidade—a propriedade que torna um fenômeno *mnêmico*—que facilita e incentiva a investigação científica (Colaço, 2022a, 2022b). Definições abrangentes como essa lançam hipóteses e meios de investigação de associações entre diversos fenômenos. No caso da hipótese da memória coletiva, para que ela seja considerada como *memória*, supõe-se que há associações importantes entre a memória coletiva e a memória individual apesar de suas diferenças. Tal suposição leva à investigação dessas associações. Além de incentivar a investigação científica, tal definição abrangente de mnemicidade, vista no Critério 2, se justifica devido ao fato de não haver uma definição a priori de memória que seja amplamente aceita e, portanto, nenhuma deve ser previamente rejeitada a menos que se

investiguem e se esclareçam as possíveis ou não associações entre os fenômenos (Colaço, 2022a). Alguém poderia argumentar que esses critérios são muito permissivos e que seria trivial que se encontrem fenômenos que os satisfaçam. No entanto, não é trivial que algo satisfaça esses critérios, especialmente porque, por se tratar de uma hipótese científica, a satisfação dos critérios requer evidência empírica.

Note que, ao seguir Colaço (2022a, 2022b) em considerar memória como um conceito abrangente, a estratégia adotada aqui é bem diferente da estratégia de outros filósofos, pois busca apenas uma pequena semelhança com outros fenômenos mnêmicos (Critério 2) ao invés de buscar alta semelhança (Michaelian, 2012, Theiner, 2013, Michaelian & Sutton, 2019, Michaelian & Perrin, 2023). Ao contrário desses filósofos, entendo que memória coletiva e memória individual devem ser fenômenos distintos (Critério 1), caso contrário, não se justificaria sua postulação. Michaelian e Perrin (2023) argumentam que, dadas as diferenças entre memória coletiva e memória individual, o termo “memória” deveria se reservar apenas à memória individual e não ser atribuído à coletividade. Diferentemente deles, sustento que existem muitos fenômenos mnêmicos distintos. Por exemplo, a memória semântica, procedural e episódica e memória de curto e longo prazo, apresentam diferenças significativas. No entanto, apesar de serem fenômenos distintos todas elas são consideradas como fenômenos mnêmicos. Da mesma forma, memória coletiva também deveria ser considerada como mnêmica apesar de ser um fenômeno distinto de memórias individuais.

Além disso, é importante ressaltar que o Critério 1 implica em que os fenômenos de memória coletiva e individual não sejam reduzíveis um ao outro pois redução implica em identidade (Piccinini, 2020). Outra forma de mostrar a não-identidade de dois fenômenos é evidenciar sua

dissociação. Por exemplo, memórias de curto prazo e longo prazo, apesar de frequentemente interdependentes, são dissociáveis. A inibição de LTP no hipocampo previne a formação de memória de curto prazo, mas não de longo prazo (Gershman, 2023). Mais ainda, a inibição de síntese de proteína no hipocampo impede a formação de memória de longo prazo, porém não a de curto prazo (Kandel, 2006).

Em suma, uma memória satisfaz o Critério 1 tornando-se *coletiva* quando: a) não se reduz/não se identifica plenamente com a memória individual; b) parte dessa impossibilidade de redução se dá porque o fenômeno tem entre uma de suas condições ser produzido por mais de uma pessoa; e c) essas pessoas devem estar em interação durante o processo de rememoração. Os motivos para a) e b) já foram discutidos nos parágrafos anteriores: para ser coletiva, a memória não pode se reduzir à (ou se identificar com a) memória individual. E para ser coletiva o motivo dessa não-redução/não-identificação deve se dar, em parte, porque uma das condições para a memória coletiva é que ela é produzida por mais de uma pessoa (um grupo, uma coletividade). No entanto, se esse grupo não estiver em colaboração, então as interações e relações causais que resultarão na memória poderão se reduzir às interações e relações causais que ocorrem dentro de cada indivíduo. Por isso, uma mera soma de pessoas não constitui uma coletividade. É necessária uma interação para que as relações causais que de fato produzem a rememoração não se reduzam aos indivíduos. Dadas essas três condições para atribuir *coletividade* a um fenômeno mnêmico, como será visto a seguir, cientistas puderam construir testes para evidenciar situações em que esses três critérios foram satisfeitos.

Um adendo final: memória coletiva vem sendo estudada em grupos pequenos e em grupos grandes (países, por exemplo) (Michaelian & Sutton, 2019). Neste texto, me atenho ao estudo de grupos pequenos,

embora seja bem possível que grupos grandes também produzam fenômenos mnêmicos, embora diferentes de fenômenos de grupos pequenos (ver Arango-Muñoz & Michaelian, 2020).

Tendo discutido os critérios a serem adotados, vamos à investigação da memória coletiva enquanto fenômeno. Em primeiro lugar, deve-se avaliar se memória coletiva satisfaz aos critérios de validação enquanto fenômeno científico (Seção 1), a saber, se ele é replicável e se é suscetível a diversas metodologias. No caso de grupos pequenos, fenômenos de memória coletiva têm sido estudados desde a década de 1980. Ren e Argote (2011) fizeram uma revisão sistemática dos estudos produzidos até então e, de fato, memória coletiva é replicável sob suas condições típicas e tem sido estudado por meio de metodologias diversas. Apresentarei uma caracterização desses fenômenos quando discutir se eles se distinguem da memória individual. No entanto, a referência mencionada fica para salientar a satisfação tanto da replicabilidade quanto da multiplicidade de métodos. Ao longo desses mais de quarenta anos de investigação, o fenômeno vem sendo replicado, tanto utilizando-se de listas de palavras, listas de desenhos, narrativas longas, testes de rememoração livre e de reconhecimento, bem como também de testes em laboratório e em contextos mais naturalistas (ver também Harris *et al.*, 2019). Tendo satisfeito os critérios para validação enquanto fenômeno científico, passemos a analisar se eles satisfazem os critérios 1 e 2⁸.

Para analisar a não-redução da memória coletiva à memória individual, psicólogos encabeçados por Wegner (1987) distinguiram dois tipos de grupos para os experimentos de memória coletiva: grupos

⁸ Ficará mais evidente a satisfação dos critérios de validação de um fenômeno científico ao discutir a satisfação dos critérios 1 e 2.

colaborativos e grupos nominais. Grupos colaborativos são formados por indivíduos que realizam as etapas de codificação (aprendizado) e retomada (teste) em conjunto, interagindo uns com os outros. Grupos nominais realizam essas etapas individualmente e o resultado deles é somado. Essa distinção tem como objetivo testar fenômenos em que as condições a) – c) se aplicam. Ao se analisar a rememoração de grupos e indivíduos, é possível ver se o conteúdo da memória do grupo se reduz à memória coletiva (condição a). Além disso, por manter outras coisas iguais, sendo uma situação controlada em laboratório, é possível inferir que a não redução da memória coletiva à memória individual se deu porque uma das condições para o fenômeno é que mais de uma pessoa estava rememorando o evento (condição b). Por fim, se as pessoas estivessem se rememorando isoladamente, então isso implicaria em que a memória coletiva se reduziria à memória individual. Por isso, a distinção entre grupos colaborativos e grupos nominais foi postulada para testar se a memória de várias pessoas se lembrando sozinha era igual ou não à memória de várias pessoas se lembrando em colaboração (condição c) (Ren & Argote, 2011). Em outras palavras, esse fenômeno emergiria apenas quando houvesse coletividade factual (interação real) e não apenas nominal.

De fato, várias pesquisas evidenciaram que, dependendo das estratégias adotadas pelo grupo colaborativo, a performance de grupos colaborativos ultrapassou significativamente a performance de grupos nominais (Ren & Argote, 2011, Harris *et al*, 2022). Com isso, psicólogos passaram a postular o fenômeno de memória coletiva. Suas propriedades estáveis envolvem a retenção e retomada de um número alto de itens em poucas (ou até mesmo uma) apresentação; com alta acurácia. Nisso já reside uma diferença entre memória coletiva e individual: além de não se reduzirem, as propriedades da memória

coletiva se distinguem das propriedades das memórias de curto e longo prazo, vistas na Seção 1. De fato, nenhum tipo de memória individual tem esse conjunto estável de propriedades que a memória coletiva exhibe (Kandel, 2006, Baddeley *et al.*, 2020).

Além disso, as condições em que as memórias coletivas emergem são diferentes das memórias individuais porque dependem da interação de vários indivíduos, algo que não ocorre, obviamente, nas memórias individuais. Mais ainda, não é qualquer tipo de interação em que as propriedades da memória coletiva, conforme vistas no parágrafo anterior, emergem. Por exemplo, se todos os membros do grupo usam a mesma estratégia para se lembrarem, há uma inibição colaborativa, em que o resultado dos grupos colaborativos, embora melhor do que a média dos resultados individuais, se torna inferior ao resultado dos grupos nominais (Theiner, 2013). Quando indivíduos ficam corrigindo uns aos outros, o resultado também é inferior (Harris *et al.*, 2019). Quando há um líder guiando o processo, os resultados são consistentes com as propriedades descritas no parágrafo anterior (Theiner, 2013). Além disso, quando os indivíduos utilizam estratégias complementares, como cada um focar nos itens ou detalhes da narrativa com os quais têm mais familiaridade, o fenômeno emerge (Harris *et al.*, 2022)⁹.

Como já mencionado, a estabilidade das propriedades da memória coletiva dentro das condições mencionadas se manteve não somente diante do critério da replicabilidade como também diante de uma multiplicidade de metodologias de investigação. Por isso, a memória coletiva satisfaz aos critérios para ser considerada um fenômeno

⁹ Na nota de rodapé 3, menciono a definição de Kaiser e Krückel (2016) de fenômeno. É interessante notar que mesmo nessa definição um pouco mais restrita, a memória coletiva satisfaz a definição, pois exige que objetos (indivíduos, nesse caso) interajam entre si de acordo com algum tipo de organização. Considerar a definição delas de fenômeno é um argumento a mais de que memória coletiva seja um fenômeno coletivo.

científico. Mais do que isso, dada sua não redução à memória individual e suas diferenças de propriedades estáveis e condições nas quais emergem, a memória coletiva satisfaz também ao Critério 1. Resta saber se esse fenômeno distinto também é mnêmico, isto é, se satisfaz também ao Critério 2.

O fato de serem usados testes de memórias de diversos tipos para estudar a memória coletiva evidencia que esse fenômeno também é mnêmico. Nesses testes, a performance depende de informações adquiridas em experiência prévia, exatamente o que é exigido a partir do Critério 2. Ou seja, eles testam se experiências prévias do grupo são usadas para modular o comportamento/performance posterior (nos testes). Além disso, como vem sendo recentemente argumentado, as interações que fazem emergir a memória coletiva e a performance da memória coletiva em si são usadas posteriormente para coordenar as ações do grupo. Por exemplo, em alguns casos, o conteúdo lembrado é usado como base para um teste de tomada de decisões em grupo. De fato, quando isso ocorre, a performance da memória coletiva é retomada e usada para conduzir a tomada de decisão em grupo (Theiner *et al*, 2010, Michaelian & Sutton, 2019).

Dessa forma, memória coletiva satisfaz os critérios de validação de um fenômeno científico, por ser replicável e suscetível de investigação sob metodologias diversas. Além disso, memória coletiva se distingue das memórias individuais (Critério 1), o que justifica sua postulação enquanto fenômeno diferente e digno de ser investigado à parte. Mais ainda, apesar de ser diferente das memórias individuais, a memória coletiva se mantém dentro de um conjunto de fenômenos mnêmicos (Critério 2). Por isso, há boas razões para considerar que memória coletiva é um fenômeno que deve fazer parte da teorização científica

sobre memória. Na próxima seção, discuto o que isso pode ou não implicar para discussões sobre consciência de grupo.

3. Memória coletiva e consciência coletiva: de que tipos?

Em um artigo provocativo intitulado “If materialism is true, the United States is probably conscious”¹⁰, Eric Schwitzgebel (2015) faz uma argumentação interessante a respeito do conteúdo deste livro e que toca neste capítulo, em especial. Basicamente, sem apresentar as minúcias do argumento, ele menciona que de acordo com o materialismo, todas as explicações têm base física. Dessa forma, a consciência, tal qual a cognição, se explica por meio de entidades/objetos reais em suas interações umas com as outras dentro de uma organização específica. Em geral, tanto a cognição quanto a consciência—ele chama de estados psicológicos, em geral—se explicam por esses critérios físicos. Ele, então, passa a apresentar diferentes animais que são considerados conscientes e têm as mais diversas formas de organização entre si, com diferentes componentes internos. O que têm em comum é uma resposta comportamental “inteligente”¹¹. Pelo fato de os Estados Unidos, como um conjunto, apresentar um comportamento “inteligente”¹² e tal comportamento ser resultado da interação de entidades organizadas de uma certa forma leva à possibilidade de que os Estados Unidos são conscientes. Usando alguns exemplos de capacidades cognitivas de grupo que não se reduzem à capacidade cognitiva de seus membros, então tal comportamento inteligente do grupo não se reduz ao comportamento dos

¹⁰ Tradução: “Se o materialismo é verdadeiro, os Estados Unidos são, provavelmente, conscientes”.

¹¹ Adaptativa talvez seja o termo mais ideal, mas ele usa a palavra “smart” e “dumb” para falar desses comportamentos.

¹² Ele reconhece que, na maior parte das vezes esse pode não ser o caso dos Estados Unidos da América.

membros do grupo e a consciência atribuída a partir desse comportamento também não se reduziria à consciência dos membros.

Apesar de não estabelecer as coisas dessa forma, Schwitzgebel dá exemplos de capacidades cognitivas de grupo para evidenciar o comportamento inteligente de grupos e, a partir disso, inferir a possibilidade de consciência coletiva. Por isso, simplificarei o argumento dele da seguinte forma: se grupos têm capacidades cognitivas enquanto grupos, então, é bem provável—ou não tem nenhum motivo forte para rejeitar essa possibilidade—de que esses grupos também tenham consciência enquanto grupo.

Pelo fato de memória coletiva ser, como argumentei na seção anterior, um fenômeno científico, podemos inferir, inspirados pelo argumento de Schwitzgebel, que consciência coletiva também o é? Dado o conceito de fenômeno apresentado na Seção 1, há duas possibilidades de fazer inferências confiáveis acerca de consciência coletiva a partir da investigação de memória coletiva enquanto fenômeno científico:

- a) se consciência for uma das *condições* de todos os fenômenos mnêmicos, então o estabelecimento de memória coletiva enquanto fenômeno implicaria na alta probabilidade de também o ser.
- b) se consciência é uma das *propriedades estáveis* de todos os fenômenos mnêmicos, então o estabelecimento de memória coletiva enquanto fenômeno implicaria na alta probabilidade de consciência coletiva também o ser.

Essas possíveis inferências confiáveis levam a conclusões probabilísticas (“implica na alta probabilidade”) e não necessárias (“implica na certeza”) porque nem sempre as condições de um

fenômeno são necessárias e quase nunca as propriedades estáveis de um fenômeno são necessárias (Boyd, 1991, 2021). Isto é, tanto as condições de um fenômeno quanto o fenômeno em si são conjuntos de propriedades que *tendem* a coocorrer, mas isso não significa que em todos as instâncias do fenômeno, todas as propriedades estarão presentes. A ciência quase nunca trabalha com condições *necessárias* no sentido filosófico de que algo seria o fato em todos os mundos possíveis (de Brigard, 2023). Por isso, se as premissas das inferências a) e b) estiverem certas, então o estabelecimento de memória coletiva apenas torna alta a probabilidade de consciência coletiva.

O problema com essas duas possibilidades é que mesmo que as premissas de a) e b) estejam corretas, isso não implica em que a consciência que é condição para ou propriedade da memória coletiva seja ela mesma coletiva. É possível que consciência individual seja uma condição para memória coletiva ou que consciência individual seja uma propriedade da memória coletiva e que daí nada implique em uma consciência coletiva. Uma possível saída para um schwitzgebeliano é postular que consciência *da entidade da qual o fenômeno mnêmico* é ou uma condição para ou uma propriedade de todos os fenômenos mnêmicos. Dessa forma, se a entidade da qual emerge a memória coletiva é um grupo colaborativo, então a consciência é de um grupo colaborativo.

Com a mudança dessas premissas surge um outro problema. Considere a), em que consciência é uma condição para fenômenos mnêmicos. Condição aqui, desde a Seção 1, se refere a condições externas ao fenômeno, que permitem que ele emergja a partir de um mecanismo (Craver & Darden, 2013, Kaiser & Krickel, 2016). A questão, no caso de a), é que condições externas não fazem parte do fenômeno em si (Craver & Darden, 2013). Em outras palavras, suas propriedades são dissociadas. Por isso, não se pode inferir de modo seguro que a

propriedade coletiva do fenômeno implique na propriedade coletiva de todas as suas condições.

No caso de b), as propriedades internas de um fenômeno podem ser do nível do fenômeno em si ou pode ser do nível dos componentes do fenômeno (níveis inferiores) (Craver, 2007, Craver & Darden, 2013). Por isso, não se pode estabelecer que, mesmo que a memória seja coletiva, a consciência, caso seja uma propriedade da memória, seja ela também coletiva, pois ela pode ser uma propriedade dos componentes do fenômeno. No caso da memória coletiva, que é composta por indivíduos, essa propriedade pode pertencer aos indivíduos e não à coletividade. Por isso, não temos nenhuma indicação de que a mudança das premissas, de que a consciência deve pertencer à entidade da qual o fenômeno mnêmico emerge, está correta. Por isso, não há bons motivos para inferir da postulação de memória coletiva enquanto fenômeno científico a postulação de consciência coletiva enquanto fenômeno científico também.

Ainda assim, um schwitzgebeliano pode insistir que consciência associada a uma capacidade coletiva sempre vem a nível do fenômeno. Por exemplo, a maioria das pessoas não diria que os neurônios estão conscientes. A consciência emerge da interação deles em processos cognitivos (Schwitzgebel, 2015). Essa analogia, no entanto, perde sua força no caso de memória coletiva porque a maioria das pessoas não diria que os seres humanos do grupo colaborativo não estão conscientes, ao contrário. Por causa da presença da consciência nos componentes do grupo, a suspeita quanto à consciência a nível superior, do grupo em si, permanece.

Schwitzgebel (2015), no entanto, muda o ônus da prova e argumenta que não há nenhum motivo válido para rejeitar essa possibilidade¹³. Todos os mais importantes argumentos apresentados contra a consciência coletiva acabariam por negar a atribuição de consciência para alguns animais mais distintos do ser humano e para possíveis alienígenas completamente diferentes do ser humano. Portanto, retomando a simplificação do argumento feita no começo desta seção, não há nenhum motivo válido para rejeitar a possibilidade de que, se grupos têm capacidades cognitivas enquanto grupos, então, grupos provavelmente têm consciência enquanto grupos. Esse argumento também é problemático diante da pergunta à qual me detenho—que se refere a fenômenos científicos—porque não ter motivos para rejeitar uma possibilidade não é motivo para que ela faça parte da teorização científica. Ou seja, o fato de que um fenômeno seja uma possibilidade, mas cujo argumento em favor dele seja uma mudança do ônus da prova, ou seja, não há evidências a favor, só não há evidências contra, não é suficiente para estabelecê-lo enquanto fenômeno científico.

Isso fica claro ao lembrar a definição de fenômeno científico apresentada na Seção 1. Fenômenos científicos são conjuntos estáveis de propriedades. Quais são as propriedades estáveis da consciência? Não há nenhuma teoria sobre consciência que esteja perto de uma unanimidade e diferentes teorias científicas sobre consciência estudam diferentes conjuntos de propriedades. Qual seria o conjunto de propriedades da consciência coletiva? Além disso, esse conjunto de propriedades é replicável? Está sujeito a investigação sob metodologias

¹³ Interessantemente, Theiner *et al* (2010) usam do mesmo argumento para defender cognição coletiva, mas rejeitam a aplicação desse argumento para defender consciência coletiva, pois o acham muito contraintuitivo. No entanto, eles reconhecem que o argumento para defender cognição coletiva também não é intuitivo. Por isso, não parece haver por que alguém que se utiliza desse argumento para defender a cognição coletiva rejeite o mesmo argumento quanto à consciência coletiva.

diversas? Apenas entender a consciência coletiva como possibilidade não é suficiente para validá-la como fenômeno científico e, até o momento, não parece haver razões para tal validação. Por isso, o argumento de Schwitzgebel não se sustenta diante da questão levantada, que se refere à postulação e validação da consciência coletiva como fenômeno científico a partir da validação da memória coletiva como fenômeno científico. E, ainda que o argumento fosse convincente, a relação entre fenômenos mnêmicos e consciência é bem complicada e deveria ser levada em conta nesse tipo de inferência.

É complicada primeiro porque, como mencionado no parágrafo anterior, diferentes teorias e projetos de pesquisa investigam conjuntos diferentes de propriedades sob a alcunha de “consciência”, que alguns chamam de dimensões de consciência. As dimensões de consciência, em geral, são inferidas a partir de capacidades cognitivas quando a expressão verbal não é uma possibilidade (Andrews, 2024). Birch e colegas (2020) apresentam cinco dimensões de consciência tem sido estudada em organismos não-linguísticos (tanto animais não-humanos quanto bebês humanos). Essas dimensões se apresentam em diferentes graus em cada organismo. As dimensões apresentadas são: riqueza perceptual; riqueza emocional; unidade sincrônica da consciência; unidade diacrônica da consciência; e autoconsciência. Não é difícil notar que certas capacidades cognitivas, como a percepção, permitem inferir graus de consciência em determinadas dimensões, como a riqueza perceptual. Por isso, validar certos fenômenos cognitivos potencialmente¹⁴ permite realizar certas inferências acerca de certas dimensões de consciência¹⁵. Então a primeira

¹⁴ “Potencialmente” é usado aqui, pois mesmo Birch *et al* (2020) argumentando nesse sentido, reconhecem que as conclusões de tais inferências devem ser alcançadas com muita cautela, levando em conta diversos fatores que fogem do escopo da argumentação deste capítulo.

¹⁵ Mais uma vez, conforme discutido nos parágrafos anteriores, no caso de capacidades cognitivas coletivas, há a possibilidade de que a consciência inferida a partir da capacidade coletiva seja

complicação é que quando se fala de consciência, deve-se esclarecer de que dimensões e em que graus se está falando.

Ao discutir memória aparece a segunda complicação, pois existem diferentes tipos de memória a partir das quais podem ser feitas diferentes inferências a respeito de diferentes dimensões de consciência. É possível, no entanto, que a maioria dos tipos de memória tem uma ou outra dimensão de consciência entre suas propriedades estáveis. O exemplo mais claro é da memória episódica e da memória semântica. Tulving (2005) considerou que na memória semântica está presente uma consciência noética e na memória episódica está presente uma consciência auto-noética. Esses tipos de consciência derivam da influência husserliana que Tulving tinha e são pouco operacionalizáveis na prática científica. Numa tentativa de tradução para as dimensões de consciência mencionadas acima, auto-noese envolveria autoconsciência e unidade diacrônica da consciência em alto grau enquanto a noese envolveria apenas a unidade sincrônica da consciência.

Contudo, alguns pesquisadores consideram que não são a essas dimensões de consciência que esses dois tipos de memória estão associados (Andonovski, 2023). Addis (2020), por exemplo, defende que o que diferencia memória semântica de episódica é o grau de riqueza perceptual e, talvez, o grau de riqueza emocional. Outros tipos de memória estariam associados a outras dimensões de consciência. Inferir consciência a partir de memória, de todo modo, depende tanto da dimensão de consciência tratada quanto do tipo de memória.

No caso da memória coletiva, alguns defendem que ela é uma memória episódica (Michaelian & Sutton, 2019, Arango-Muñoz &

consciência individual. A discussão sobre a relação entre memória e consciência coletivas surge, como visto, a partir da suposição de que o argumento de Schwitzgebel é convincente.

Michaelian, 2020). Nesse caso, dependendo do que se entende por memória episódica, pode-se inferir consciência coletiva de uma certa dimensão (ou autoconsciência ou riqueza perceptual, conforme discutido nos dois parágrafos anteriores). No entanto, essa conclusão de que memória coletiva é memória episódica desconsidera aspectos cruciais da memória episódica que não estão presentes na memória coletiva, como as diversas formas de construção, tanto na codificação quanto na rememoração, como Michaelian e Perrin (2023) admitem posteriormente. De fato, levando em conta que o fenômeno de memória coletiva é dissociado da memória individual, como visto na Seção 2, então as inferências que ambos sustentam devem ser diferentes. Por isso, ainda que se aceite que memória coletiva implica em consciência coletiva, fica em aberto que dimensão está sendo inferida.

Em suma, inferir que consciência coletiva deve ser um fenômeno científico porque memória coletiva, enquanto capacidade cognitiva, é um fenômeno científico é uma inferência não confiável, pois não aponta as propriedades estáveis da consciência coletiva nem implica na replicação e na suscetibilidade de investigação sob metodologias diversas dessas propriedades estáveis. Além disso, pelo fato de a memória coletiva e da memória individual emergirem de componentes bem diferentes (os componentes da memória individual em si não são conscientes enquanto os componentes da memória coletiva são) enfraquece a associação entre consciência e capacidade cognitiva em grupos. Apesar disso, caso se aceite que memória coletiva implica em consciência coletiva, essa inferência é bastante complicada, pois envolve diferentes dimensões de consciência e diferentes tipos de memória associadas a essas dimensões.

Conclusão

Neste capítulo, discuti memória coletiva e consciência coletiva sob uma questão mais fraca do que normalmente vem sendo discutida. Digo mais fraca, pois não abordo o problema de um ponto de vista metafísico, como os demais capítulos se propõem a tratar. Antes, minha abordagem à questão da memória e consciência coletiva se reduzem plenamente a questões de prática científica. Em outras palavras, no lugar de questionar se tais fenômenos são fenômenos reais/existentes, questionei se tais fenômenos podem se classificar como científicos. Após apresentar a definição de fenômeno científico a ser abordada, defendi que memória coletiva é um fenômeno científico. Também argumentei que tal conclusão não permite uma inferência confiável de que consciência coletiva também é um fenômeno científico. Essa argumentação, no entanto, depende das evidências disponíveis até o momento. Pode ser que achados futuros comprovem a validade das conclusões aqui chegadas ou que provem que elas estão erradas. Além disso, é importante ressaltar que validar memória coletiva como fenômeno científico não implica em considerá-la como real e que rejeitar consciência coletiva como fenômeno científico não implica em considerá-la como não existente, como discutido ao fim da Seção 1.

Referências

- ADDIS, D. R. Mental Time Travel? A Neurocognitive Model of Event Simulation. *Review of Philosophy and Psychology*, v. 11, n. 2, p. 233–259, 2020.
- ANDONOVSKI, N. Autooesis and the Galilean science of memory: Explanation, idealization, and the role of crucial data. *European Journal for Philosophy of Science*, v. 13, n. 3, p. 42, 2023.

- ANDREWS, K. "All animals are conscious": Shifting the null hypothesis in consciousness science. *Mind & Language*, p. mila.12498, 2024.
- ARANGO-MUÑOZ, S.; MICHAELIAN, K. From Collective Memory... to Collective Metamemory? Em: FIEBICH, A. (Ed.). *Minimal Cooperation and Shared Agency*. Studies in the Philosophy of Sociality. Vol. 11. Cham: Springer International Publishing, 2020. p. 195–217.
- BADDELEY, A. D.; EYSENCK, M. W.; ANDERSON, M. C. *Memory*. 3. ed. New York: Routledge, Taylor & Francis Group, 2020.
- BECHTEL, W. *Mental mechanisms: philosophical perspectives on cognitive neuroscience*. Hove: Psychology - L. Erlbaum, 2008.
- BIRCH, J.; SCHNELL, A. K.; CLAYTON, N. S. Dimensions of Animal Consciousness. *Trends in Cognitive Sciences*, v. 24, n. 10, p. 789–801, 2020.
- BOGEN, J.; WOODWARD, J. Saving the Phenomena. *The Philosophical Review*, v. 97, n. 3, p. 303, 1988.
- BOYD, R. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, v. 61, n. 1–2, p. 127–148, fev. 1991.
- BOYD, R. Rethinking natural kinds, reference and truth: towards more correspondence with reality, not less. *Synthese*, v. 198, n. S12, p. 2863–2903, jun. 2021.
- BRIGARD, F. D. *Memory and Remembering*. Cambridge: Cambridge University Press, 2023.
- COLAÇO, D. Rip it up and start again: The rejection of a characterization of a phenomenon. *Studies in History and Philosophy of Science Part A*, v. 72, p. 32–40, 2018.
- COLAÇO, D. Recharacterizing scientific phenomena. *European Journal for Philosophy of Science*, v. 10, n. 2, p. 14, 2020.
- COLAÇO, D. What Counts as a Memory? Definitions, Hypotheses, and "Kinding in Progress". *Philosophy of Science*, v. 89, n. 1, p. 89–106, 2022a.
- COLAÇO, D. Why studying plant cognition is valuable, even if plants aren't cognitive. *Synthese*, v. 200, n. 6, p. 453, 2022b.

- CRAVER, C. F. *Explaining the brain: mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon press, 2007.
- CRAVER, C. F.; DARDEN, L. *In search of mechanisms: discoveries across the life sciences*. Chicago: University of Chicago Press, 2013.
- GERSHMAN, S. J. The molecular memory code and synaptic plasticity: A synthesis. *Biosystems*, v. 224, p. 104825, 2023.
- GLENNAN, S. *The new mechanical philosophy*. Oxford: Oxford university press, 2017.
- HARINEN, T. Mutual manipulability and causal inbetweenness. *Synthese*, v. 195, n. 1, p. 35–54, 2018.
- HARRIS, C. B. et al. Features of Successful and Unsuccessful Collaborative Memory Conversations in Long-Married Couples. *Topics in Cognitive Science*, v. 11, n. 4, p. 668–686, 2019.
- HARRIS, C. B. et al. Ageing Together: Interdependence in the Memory Compensation Strategies of Long-Married Older Couples. *Frontiers in Psychology*, v. 13, p. 854051, 2022.
- KAISER, M. I.; KRICKEL, B. The Metaphysics of Constitutive Mechanistic Phenomena. *The British Journal for the Philosophy of Science*, v. 68, n. 3, p. 745–779, 1 set. 2017.
- KANDEL, E. R. *In search of memory: the emergence of a new science of mind*. New York: Norton, 2006.
- MICHAELIAN, K. Is external memory memory? Biological memory and extended mind. *Consciousness and Cognition*, v. 21, n. 3, p. 1154–1165, 2012.
- MICHAELIAN, K.; PERRIN, D. La métaphysique de la mémoire collective. Em: SOUCHAY, C.; LUCIANI, I. (Eds.). *La mémoire à l'épreuve de l'interdisciplinarité : sciences humaines et cognitives*. Corps & âmes. Aix-en-Provence: Presses Universitaires de Provence, 2023.
- MICHAELIAN, K.; SUTTON, J. Collective mental time travel: remembering the past and imagining the future together. *Synthese*, v. 196, n. 12, p. 4933–4960, 2019.

PICCININI, G. *Neurocognitive mechanisms: explaining biological cognition*. First edition ed. Oxford: Oxford University Press, 2020.

POTOCHNIK, A. *Idealization and the aims of science*. Chicago: The University of Chicago Press, 2017.

REN, Y.; ARGOTE, L. Transactive Memory Systems 1985–2010: An Integrative Framework of Key Dimensions, Antecedents, and Consequences. *Academy of Management Annals*, v. 5, n. 1, p. 189–229, 2011.

SCHWITZGEBEL, E. If materialism is true, the United States is probably conscious. *Philosophical Studies*, v. 172, n. 7, p. 1697–1721, 2015.

THEINER, G. Transactive Memory Systems: A Mechanistic Analysis of Emergent Group Memory. *Review of Philosophy and Psychology*, v. 4, n. 1, p. 65–89, mar. 2013.

THEINER, G.; ALLEN, C.; GOLDSTONE, R. L. Recognizing group cognition. *Cognitive Systems Research*, v. 11, n. 4, p. 378–395, 2010.

TULVING, E. Episodic memory and autoeogenesis: uniquely human? Em: TERRACE, H.; METCALFE, J. (eds.). *The missing link in cognition: origins of self-reflective consciousness*. Oxford: Oxford University Press, 2005.

WEGNER, D.M. Transactive memory: A contemporary analysis of the group mind. Em MULLEN, B.; GOETHALS, G.R. (Eds.). *Theories of group behavior*. Vol. 9. New York: Springer, 1987, p. 185–208.

WOODWARD, J. *Making things happen: a theory of causal explanation*. Oxford: Oxford University Press, 2003.

6

ONTOLOGIA SOCIAL PARA INSETOS EUSSOCIAIS

Valdenor Monteiro Brito Júnior

A ontologia social contemporânea busca investigar a estrutura metafísica do mundo social em seus diferentes componentes, bem como em suas interrelações a outros componentes da realidade. Investiga a natureza de entidades sociais, como os tipos sociais (casamento, dinheiro, etc.) e os grupos sociais (clubes, empresas, etc.).

Dois componentes de grande monta na maneira como a ontologia social é articulada diz respeito a um mentalismo e a um antropocentrismo: o mundo social depende da mente humana, esta entendida em termos de estados intencionais e/ou atitudes proposicionais como crença, intenção prévia, intenção em ação, etc. A linguagem humana também recebe destaque teórico, por exemplo, na maneira como atos de fala declarativos podem criar estados de coisas apenas por declará-los como existindo.

A questão da ontologia social entre os animais não-humanos é ainda pouco tratada na literatura. Ainda que se reconheçam certos fatos sociais a respeito dos animais não-humanos, a ontologia social contemporânea é articulada de maneira a tomar o caso humano por paradigma sob a qual entender o social. Dessa forma, o caso dos animais não-humanos se torna, na melhor das hipóteses, dependente de seu encaixe em requisitos formulados tendo animais humanos em mente, ou, na pior das hipóteses, um caso não-genuíno do fenômeno social. Abordarei o tratamento de animais não-humanos na obra de John Searle como exemplo de como essa questão é tratada por um dos pioneiros da área.

O objetivo do presente artigo é contrapor-se ao cenário atual, buscando pensar a ontologia social de forma não-antropocêntrica, mas sim biocentrada, entendendo o fenômeno social à luz de sua existência ao longo da evolução dos seres vivos em nosso planeta, com especial destaque à evolução dos animais.

Para tanto aqui destacarei o caso dos insetos eussociais, como o de várias espécies de formigas e abelhas, defendendo que tais seres também possuem ontologia social (específica à espécie pertinente) com tipos sociais, em um sentido em par com o dos seres humanos. Formulada uma defesa desse ponto para o caso das formigas e abelhas, entendo que o caso se generalize para todos os outros insetos eussociais e mesmo para todos os outros animais eussociais, vertebrados e invertebrados.

1. Animais Não-Humanos em Ontologia Social

Para ilustrar a maneira insatisfatória com que a questão da sociabilidade dos animais não-humanos é desenvolvida na ontologia social contemporânea, tomarei por exemplo a obra de John Searle, tendo em vista seu pioneirismo na área. Enquanto o Programa Intencionalista Forte (pautado na necessidade de intencionalidade coletiva para todos os tipos sociais) esposado por Searle tenha sido muito criticado na literatura especializada, e, portanto, esteja longe de ser uma unanimidade, a maneira como os animais não-humanos são retratados no contexto de sua abordagem de ontologia social não foi objeto de contestação significativa, de maneira que ainda se trata de uma premissa no background do debate contemporâneo que não tem sido criticamente examinada.

Em sua primeira abordagem para a Ontologia Social, em “*The Construction of Social Reality*” (1995), Searle chega a abordar a questão de fatos sociais em animais não-humanos e mesmo de seu ter intencionalidade coletiva. Em um quadro esquemático, Searle descreve o seguinte caminho: há fatos físicos brutos e fatos mentais. Entre os fatos mentais, alguns são não-intencionais, como o sentir dor, e outros são intencionais, como um desejo ou uma crença. Entre esses fatos mentais intencionais, há os singulares e os coletivos. Searle aqui equaliza fatos mentais intencionais coletivos com fatos sociais, usando o exemplo de hienas caçando um leão. Então, Searle entende que – ao menos hipoteticamente – certos animais possuam estados mentais de intencionalidade coletiva que, para Searle, seriam necessários à cooperação/sociabilidade genuína. Contudo, não estamos aqui ainda diante do objeto próprio da Ontologia Social, e não há tipos sociais. Entre esses fatos mentais intencionais coletivos, há uma categoria especial deles que diz respeito à atribuição de função para fatos físicos brutos. Essa atribuição pode ser de uma função agentiva ou não-agentiva. E, na agentiva, pode ser causal ou de status. É apenas nas funções de status (como um tipo de função agentiva) atribuídas a fatos físicos brutos por intermédio de fatos mentais intencionais coletivos que chegamos no objeto próprio da Ontologia Social, a que Searle denomina de fatos institucionais. Assim, os animais não-humanos ficam excluídos da Ontologia Social, mesmo quando se trate de animais sociais, porque não fariam atribuição coletiva de funções de status.

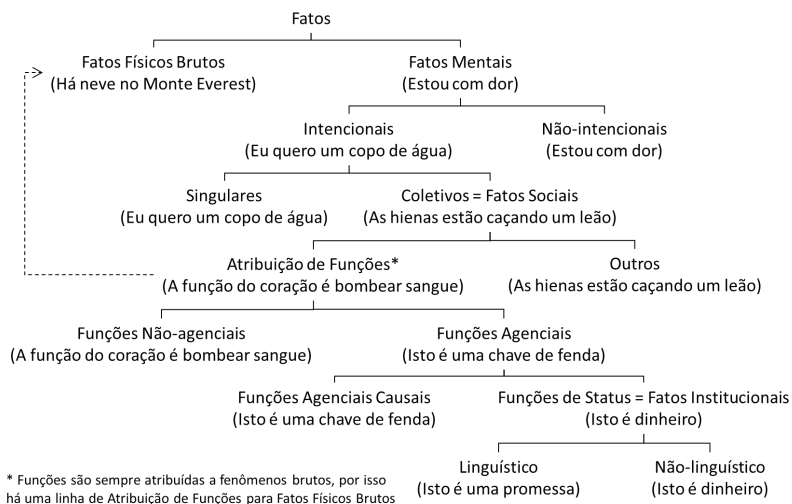


Figura 1. Hierarquia taxonômica de (certos tipos de) fatos. Searle, 1995, p. 121.

Em sua formulação mais recente, na obra *“Making the Social World”* (2010), com mudanças significativas em relação à anterior, a questão dos animais sociais não-humanos continua relegada à irrelevância teórica. Searle agora dá destaque aos atos de fala. Ele faz uma distinção entre estados mentais intencionais e relaciona-os aos atos de fala. Atos de fala declarativos, ou Declarações, permitem fazer coisas existirem por meio de sua representação como existindo, por declará-las como existentes. Nenhum estado mental intencional consegue fazer isso e, portanto, os atos de fala declarativos não possuem um atributo que falta no mundo mental pré-linguístico, sendo, assim, uma novidade introduzida pela linguagem.

Essa nova formulação acarreta problemas ainda mais significativos ao caso dos animais não-humanos em ontologia social: nenhum animal não-humano possui linguagem da maneira como a possuímos e, mesmo que se tente adaptar o ponto de Searle de atos de fala à comunicação animal ou comportamento verbal animal, parece

difícil dizer que animais tenham atos de fala declarativos. Então, a ontologia social de Searle não contempla ontologia social para seres não-linguísticos, mesmo que se trate de seres com estados mentais intencionais coletivos.

Atos de fala declarativos parecem o suprassumo da distintividade humana, tornada possível por meio de uma linguagem simbólica representacional que emergiu em nossa espécie, o *Homo sapiens*, durante a revolução simbólica cuja ocorrência estima-se ter sido em torno de 80.000 anos atrás. Note que antes da revolução simbólica já existiam *Homo sapiens*, então nem mesmo o *Homo sapiens* tinha ontologia social antes desse momento, quanto mais os animais não-humanos! E nem outros homínídeos ancestrais já extintos do clado em que estamos inseridos teriam ontologia social: nem australopithecíneos, nem *Homo erectus*, nem *Homo habilis*. O critério de Searle para ter ontologia social é muito rigoroso.

Muitas partes da abordagem de Searle foram criticadas. Por exemplo, a necessidade da intencionalidade coletiva é criticada na literatura¹. Também a maneira como a abordagem de Searle parece desconsiderar que nem todos os tipos sociais são constituídos da mesma maneira. Khalidi (2015) em seu artigo sobre as três categorias de tipos sociais defende que a abordagem de Searle apenas é acurada para os tipos sociais que dependem de atitudes intencionais tanto para o tipo como para suas instâncias concretas. Contudo, se assim desmontarmos a abordagem de Searle, podemos chegar a novas conclusões sobre a ontologia social dos animais não-humanos. É o que faremos a seguir, tomando por exemplo os insetos eussociais, mais especificamente das colônias de formigas e abelhas.

¹ Para uma revisão da literatura respectiva, veja o 1º capítulo de minha tese de doutorado (Brito Jr., 2021).

2. Socialidade e Eussocialidade em Zoologia

Para indagarmos acerca da ontologia social em insetos eussociais, algumas distinções precisam ser feitas no que diz respeito à pesquisa biológica na seara do comportamento social².

Um primeiro conceito seria o de socialidade. Wilson (1975, p. 7) define sociedade, em termos da biologia comportamental, como um grupo de indivíduos pertencentes à mesma espécie e organizados de uma maneira cooperativa. Ele afirma que as definições de sociedade e do social em sociobiologia precisam ser as mais abrangentes possíveis, não excluindo nenhum fenômeno interessante intuitivamente visto como social. De fato, comunicação recíproca de natureza cooperativa, transcendendo mera atividade sexual, é o critério intuitivo essencial de uma sociedade (Wilson, 1975, p. 7). Note que Wilson explicitamente estabelece o caráter intraespecífico do social, em termos de atividade cooperativa. Esta é importante porque a formação de grupo³ não é suficiente para vida social genuína. A mera agregação de organismos⁴, elicitada por estímulo mutuamente atrativo, não forma sociedades animais. Wilson considera que, enquanto agregação, comportamento sexual e proximidade física em área territorial compartilhada não são suficientes para estabelecer sociedades, esses fenômenos são propriedades presentes em sociedades verdadeiras. Isto é, podemos

² Parte do material a seguir foi aproveitado da minha tese de doutorado "Ontologia Social Não-Mentalista: uma abordagem biocentrada" (2021; p. 306-308).

³ "Group: a set of organisms belonging to the same species that remain together for any period of time while interacting with one another to a much greater degree than with other conspecific organisms. The word group is thus used with the greatest flexibility to designate any aggregation or kind of society or subset of a society." (Wilson, 1975, p. 8)

⁴ "Aggregation: a group of individuals of the same species, comprised of more than just a mated pair or a family, gathered in the same place but not internally organized or engaged in cooperative behavior. Winter congregations of rattlesnakes and ladybird beetles, for example, may provide superior protection for their members, but unless they are organized by some behavior other than mutual attraction they are better classified as aggregations than as true societies." (Wilson, 1975, p. 8)

dizer que são derivativamente sociais se forem feitas em grupos sociais ou por indivíduos de uma espécie social.

Eusociabilidade (cunhado por Batra, 1968) é definida por Wilson e Hölldobler (2005) como um nível evolutivamente avançado de existência colonial onde membros adultos de uma colônia animal pertencem a duas ou mais gerações sobrepostas, cuidam cooperativamente dos infantes entre eles, e são divididos em castas reprodutivas e não-reprodutivas (ou menos reprodutivas). Nessa mesma direção, Nowak, Wilson e Tarnita (2010, p. 1057) entendem “eussocialidade” como a característica de espécies sociais onde membros adultos compõem castas reprodutivas e parcialmente não-reprodutivas, sendo as últimas ocupadas com o cuidado dos jovens. Isso significa que a eussocialidade envolve cuidado cooperativo (inclusive de prole não aparentada diretamente), gerações sobrepostas e divisão do trabalho entre essas diferentes castas. Em razão disso, a eussocialidade é um fenômeno mais estrito e muito raro na natureza. Insetos sociais são o melhor exemplo desse conceito, onde a sociedade é dividida entre a rainha e machos reprodutivos responsáveis pela reprodução, de um lado, e soldados e operárias não-reprodutivos que tornam possível a vida colonial, de outro lado⁵.

A tabela abaixo apresenta os vários graus de sociabilidade entre espécies de insetos, na formulação clássica em Wilson (1975), sendo o eusocial o nível mais elevado possível:

⁵ Há algumas espécies e gêneros de formigas nos quais a reprodução não é realizada por uma rainha, mas por meio de uma operária com ovários desenvolvidos e espermateca funcional, denominada “gamergate” (Peeters & Crewe, 1984; Peeters, 1991). É o caso da *Dinoponera lucida*, endêmica na Mata Atlântica brasileira (Peixoto et al, 2008). Mas por ser um caso raro, sempre falaremos no texto em termos da rainha sendo a reprodutora.

Graus de Sociabilidade	Qualidades de Sociabilidade		
	Cuidado cooperativo da prole	Castas Reprodutivas	Sobreposição de gerações
<i>Sequência Parassocial</i>			
Solitária	-	-	-
Comunal	-	-	-
Quasissocial	+	-	-
Semissocial	+	+	-
Eussocial	+	+	+
<i>Sequência Subsocial</i>			
Solitária	-	-	-
Primitivamente Subsocial	-	-	-
Intermediária Subsocial I	-	-	+
Intermediária Subsocial II	+	-	+
Eussocial	+	+	+

Tabela 1. Graus de Sociabilidade nos Insetos. Wilson, 1975, p. 398.⁶

Alguns cientistas entendem que os seres humanos poderiam ser considerados eussociais, com base na presença universal de avós pós-menopausa proporcionando suporte à reprodução de sua prole nas sociedades humanas. Sob essa visão, a perda da capacidade reprodutiva – variável fisiológica – associada ao envelhecimento nas fêmeas da espécie humana estaria funcionalmente relacionada ao suporte fornecido – variável comportamental – de forma análoga a uma casta não-reprodutiva (Foster; Ratnieks, 2005; Wilson; Nowak, 2014; veja também uma consequência dessa atribuição para a questão da psicologia do suicídio em termos de um desarranjo da dimensão de autossacrifício no contexto eussocial em Joiner et al, 2016; Chiurliza et al, 2018). Contudo, tal atribuição de eussocialidade aos humanos é controvertida e não universalmente aceita, em razão da dificuldade de enquadrar a espécie humana nos termos mais estritos da eussocialidade, que demandaria uma divisão extrema do trabalho reprodutivo, também conhecida como

⁶ Para uma exposição mais recente acerca desses graus de sociabilidade, consulte Gullan; Cranston, 2005, p. 299-326.

reproductive skew – e assim remanesce controverso se a questão relativa à menopausa anteriormente mencionada seria suficiente para nos qualificar sob esse critério.

Exemplos consensuais de mamíferos eussociais são o Rato-toupeira-pelado (*Heterocephalus glaber*) e o Rato-Toupeira de Damaraland (*Fukomys damarensis*), que organizam colônias em termos de alguns poucos membros reprodutivos associados a uma maioria não-reprodutiva. Aqui focaremos a exposição nos insetos eussociais.

Assim, nesse primeiro momento, observamos que a sociabilidade dos insetos eussociais (e de outros animais eussociais de modo geral) se desenvolve de forma diferente em relação à sociabilidade humana, uma vez que entre os animais humanos não ocorre uma supressão reprodutiva da maioria do grupo social. Outra questão é que nas espécies de insetos eussociais a relação genética (*relatedness*) entre os membros cooperativos costuma ser muito alta (essa afirmação será qualificada mais adiante), enquanto no *Homo sapiens* há cooperação em larga escala entre estranhos não-relacionados. É em razão deste último ponto que alguns autores cunharam a expressão “ultrasocialidade” para o caso humano, definida por Turchin (2013, p. 62 e 65, mencionando Campbell, 1983) como a habilidade dos seres humanos em cooperar em grandes grupos (incluindo mesmo milhares e milhões de pessoas) compostos por indivíduos que não sejam proximamente relacionados em termos de genética (isto é, que não são parentes entre si), circunstâncias estas que seriam *prima facie* mais propícias para a adoção de uma estratégia comportamental egoísta ao invés de cooperativa, dado que os indivíduos não-parentes estariam em conflito direto pelos recursos.⁷ Esse conceito enfatiza a capacidade

⁷ Uma exceção seria a discussão clássica da seleção de grupo, no qual a competição entre grupos poderia favorecer a cooperação dentro do grupo caso a pressão seletiva competitiva intergruppal conseguisse superar a intragrupal, mas não entrarei nessa complexidade aqui.

humana para cooperar com muitos estranhos não-relacionados em grupos grandes.

Contudo, apesar desses diferentes caminhos de sociabilidade, tanto a humana com sua ultrasocialidade como a dos insetos eussociais com sua eussocialidade se destacam por estarem entre os níveis mais complexos conhecidos de organização social encontrados na natureza viva. Isso não deve ser entendido numa falsa aceção que associe a evolução biológica a uma ideia de “progresso” (muito menos de um progresso inevitável). A socialidade enquanto fenômeno biológico não precisa tender à eussocialidade nem à ultrasocialidade (e nem mesmo *in extremis* à criação de um novo organismo, como ocorrido na transição da unicelularidade para a multicelularidade, a qual pode ser entendida como precedida por eussocialidade entre linhagens celulares; Birch, 2017, cap. 7). Contudo, é nelas que vemos o fenômeno social de maneira mais complexa, envolvendo uma escala de cooperação entre seres vivos bem rara na natureza.

3. Ontologia Social para Insetos Eussociais

A Ontologia Social volta-se ao estudo do mundo social em termos de categorias metafísicas e seus inter-relacionamentos ontológicos com diferentes dimensões da realidade. Contudo, o paradigma de mundo social seria o humano (e quase que exclusivamente o humano, como revisado em seção anterior).

Meu argumento no presente artigo é que, se afirmarmos uma ontologia social substantiva para o caso humano, não estamos justificados em negar isso para o caso de animais eussociais, inclusive por referência aos insetos eussociais. Portanto, aceitar a legitimidade da discussão ontológica do mundo social humano, também legitima a

discussão ontológica do mundo social em animais eussociais, inclusive invertebrados.⁸

Quando consideramos possíveis motivos pelos quais afirmaríamos a ontologia social para a espécie humana, enquanto a negaríamos para espécies de insetos eussociais (com destaque para formigas, abelhas, vespas e cupins), dificilmente encontramos quaisquer bases muito claras que motivassem tal tratamento diferenciado.

Consideremos a característica presente em muitas formigas de formar trilhas químicas em sua atividade forrageadora. Também poderíamos considerar o caso de trilhas químicas aéreas deixadas por certas espécies de abelhas. Parece-me que estamos justificados em afirmar, com base no atual estado da arte em zoologia, que “trilhas químicas fazem parte da organização social de certas espécies eussociais de formigas e abelhas”. Dificilmente encontraríamos razões internas à biologia que recusassem fidedignidade a tal afirmativa. Agora note a similitude com o caso humano: “rotas de circulação de pessoas e bens fazem parte da organização social de sociedades humanas ao redor do globo”. Por “rotas de circulação de pessoas e bens” entendemos aqui itinerários de via terrestre/marítima/aérea utilizada por uma sociedade humana, em deslocamentos de um ponto a outro, onde referida rota esteja “sinalizada” para uso compartilhado: isso permite englobar desde trilhas num matagal fechado, com pouquíssima alteração do ambiente circundante, usadas para caça ou coleta conjuntas em sociedades de pequena escala, até os mais complexos sistemas de tráfego rodoviário e

⁸ De fato, entendo como legítima a abordagem ontológica para o mundo social de todos os animais (quicá organismos) sociais, mesmo daqueles que não são eussociais nem ultrasociais nos sentidos acima expostos. Contudo, neste artigo me restringirei a uma defesa dessa tese para o caso dos animais eussociais. Consulte Brito Júnior (2021) para apontamentos em torno da ontologia social para quaisquer animais (ou organismos) sociais.

mesmo aéreo implantados em sociedades industriais modernas de grande escala.

Ao olhar para o caso humano, parece natural em ontologia social afirmar que as práticas sociais envolvendo rotas sinalizadas (com diferentes características entre diferentes sociedades no tempo e no espaço) perfazem entidades sociais, tipos sociais, que podem ser analisados por intermédio do instrumental da metafísica analítica contemporânea. Esse apontamento funciona mesmo se quisermos dizer que as rotas sinalizadas são partes próprias de tipos sociais ou participam na sua constituição ao invés de serem os tipos sociais em si mesmos, mas, para fins de simplicidade na discussão a seguir, as tratarei como entidades sociais completas. Também não deve se confundir aqui a noção de “rota de circulação de pessoas e bens” com os artefatos usados para percorrer essas rotas (aviões, navios, veículos, e até mesmo estradas asfaltadas).

Dois exemplos mais específicos podem ser fornecidos para ilustrar melhor a comparação a ser feita aqui. Primeiro, o leitor pode pensar no caso de rotas comerciais. Algumas foram de enorme importância histórica, como a Rota da Seda. Tais rotas capacitam a circulação de pessoas e bens através de trocas comerciais e seriam impossíveis sem algum tipo de comunicação entre vários seres humanos perfazendo a Rota. Mas muitas delas surgem de forma espontânea através de uma multidão de interações descentralizadas, gerando aquilo que o teórico social Hayek (1967) denominou de “ordem espontânea” (Gedeon, 2015). A desvantagem desse exemplo é que pode ser objetado que rotas comerciais dependem da existência de outro tipo social tal como “dinheiro”, que nem sempre existiu nas sociedades humanas. Para não entrar nesse debate por não ser o foco aqui, também introduzo um segundo exemplo para ajudar a entender o paralelo. Sociedades caçadoras-coletoras são conhecidas por

praticarem forrageamento de local central, que é comparável à prática de formigas eussociais em forragearem a partir do ninho da colônia ou e das abelhas eussociais a partir da colmeia. Por intermédio desse tipo de forrageamento, surgem redes de interação e rotas de circulação que, via atividades comunicativas, promovem a transmissão da informação ao longo da rede (Garg et al, 2021). A adoção desse sistema de forrageamento nos distingue de hominídeos mais ancestrais e dos grandes primatas antropóides não-humanos (Garg et al, 2021; Layton; O'hara, 2010; Jang et al, 2019).

Então, meu ponto é: se entendermos que tais rotas sinalizadas de circulação de pessoas e bens são tipos sociais no caso humano, teríamos alguma justificativa para não entendermos as trilhas químicas terrestres de formigas eussociais ou as trilhas químicas aéreas de abelhas eussociais como tipos sociais das espécies pertinentes?

Uma primeira dúvida que poderia surgir é se deveríamos falar das trilhas químicas em insetos eussociais como fatos sociais ou como fatos meramente biológicos. Afinal, a capacidade para produzir os feromônios necessários à constituição dessas trilhas não é simplesmente uma parte da morfologia e fisiologia das espécies respectivas? Contudo, é importante observar que não estou aduzindo que a produção de feromônios em si seja uma marca do social ou perfaça um fato social em si mesmo. Insetos solitários também produzem feromônios. Por exemplo, para fins de acasalamento (Yan; Liebig, 2021, p. 472; Blomquist; Bagnères, 2010). De fato, os insetos eussociais e os insetos solitários compartilham em grande medida os genes essenciais a regular o desenvolvimento e a atividade neuronal subjacente ao uso de sinais químicos, mas certos genes foram cooptados para regular comunicação social e comportamento social nos eussociais (Yan; Liebig, 2021, p. 477; Amdam et al. 2006; Robinson et al. 2008), explicando assim sua diferença marcante.

Daí a escolha pelo exemplo das trilhas químicas nos insetos eussociais: as trilhas químicas são produtos de cooperação e comunicação entre as formigas/abelhas pertinentes. Vamos usar o exemplo das formigas. A trilha química formada no curso das atividades forrageadoras é constituída pela emissão de feromônios no sentido de atrair mais formigas da mesma colônia na obtenção de recursos, uma vez que os semioquímicos (ou infoquímicos) nesse caso são usados como “sinais”, uma vez que são produzidos por um emissor que tem sido moldado pela evolução para transmitir uma mensagem específica ao receptor pretendido (Chapman, 2013, p. 858). Portanto, essas trilhas químicas pressupõem comunicação entre os insetos individuais pertinentes na satisfação cooperativa de um propósito compartilhado. Sem a cooperação social subjacente, tais trilhas químicas não existiriam ou não seriam o fenômeno que são. Por exemplo, se uma aparente trilha química emergisse simplesmente como subproduto acidental da produção de feromônios em espécies solitárias e da movimentação dos indivíduos dessas espécies em um espaço comprimido, não estaríamos diante do mesmo fenômeno que a trilha química como moldada pela seleção natural para fins de cooperação eficiente numa atividade forrageadora cujo resultado beneficia a colônia. Note que a questão não é terminológica, sobre o uso do termo “trilha química”. E sim que, independentemente do nome usado, no primeiro caso estaríamos diante de certo fenômeno e no segundo caso estamos diante de outro fenômeno, sendo que este último pressupõe sociabilidade e cooperação.

Uma analogia que pode ser feita aqui para esclarecimento da questão seria a capacidade de emitir ruídos sonoros. Várias espécies animais solitárias têm capacidade de fazê-lo. Nesse sentido, trata-se de fatos a respeito de suas fisiologias e morfologias. Contudo, esse fato fisiológico e morfológico (meramente biológico) pode ser cooptado na

evolução de certas espécies para ser utilizado no decurso de atividades cooperativas, como no próprio caso dos primatas humanos, onde as cordas vocais servem para comunicação simbólica complexa.

Por outro lado, o objetor ainda poderia tentar insistir nessa questão de biológico versus social, arguindo o seguinte: no caso dos insetos eussociais, estamos diante de um comportamento sujeito a forte controle genético e inato, enquanto, no caso dos primatas humanos, estamos diante de comportamentos em grande medida aprendidos e mesmo deliberadamente escolhidos. O peso da genética no comportamento eusocial das colônias de formigas ou abelhas faria das “trilhas químicas” entidades meramente biológicas, enquanto os exemplos análogos entre os seres humanos, de “rotas de circulação de pessoas e bens”, não estariam submetidos a tal controle genético e assim estariam qualificados a serem entidades sociais.

O problema com essa objeção diz respeito a uma visão simplista da relação entre genes e comportamento, e da distinção inato/aprendido. Não revisarei aqui essa discussão em toda a sua extensão e profundidade, mas me aterei ao seguinte aspecto. O mapa das rotas de circulação de pessoas e bens desenvolvidas por uma sociedade humana específica não está “prevista” no nosso código genético, afinal, a forma que tais trilhas tomarão depende de elementos externos, desde o relevo geográfico até as cadeias produtivas de determinada sociedade. Ninguém poderia olhar para o Projeto Genoma Humano e daí inferir um mapa detalhado das redes de tráfego no Brasil no século XXI. Contudo, o mesmo ocorre para os insetos eussociais!

Formigas eussociais tenderão a produzir trilhas químicas na medida em que, por razões genéticas, têm a capacidade fisiológica e a disposição comportamental de emitir feromônios para fins comunicativos em atividades voltadas à obtenção de alimentos, tal como nós, também por

razões genéticas, tenderemos a produzir rotas de circulação de pessoas e bens por termos a capacidade fisiológica e a disposição comportamental de emitir ondas sonoras vocais no intuito de cooperarmos na obtenção de alimento (não só deste, mas o mesmo vale para os insetos eussociais, a depender da espécie as trilhas químicas podem ser usadas também para fins outros que não a obtenção de alimento, como quando usadas por formigas ocupando o papel de “patrulheira”).

Como serão exatamente essas trilhas químicas? Qual seu tamanho, quantos membros da colônia envolverão etc.? Tudo isso vai depender das circunstâncias do ambiente e do histórico prévio. Deborah Gordon (2014; 2023; Gordon et al, 2011) descreve bem como as trilhas emergem a partir de interações locais entre as formigas de uma colônia, não estando sujeitas a um controle central (e aqui nem mesmo a um “como se controle central” dos genes subjacentes).

Mesmo a noção de casta na divisão de trabalho colonial não implica em um determinismo genético linear, uma vez que, tomando formigas como referência: 1) a diferenciação de casta nas formigas é amplamente trofotênica – isto é, determinada pela dieta –, envolvendo uma alocação enviesada do volume e qualidade da comida dada às larvas, de modo que uma larva aleatoriamente selecionada pode se desenvolver para virar rainha ou operária a depender da dieta (Gullan; Cranston, 2005, p. 312); 2) em determinadas espécies, usando de exemplo estudos relativos à *Pogonomyrmex barbatus*, uma formiga que não atua como forrageadora pode mudar de papel para atuar como forrageadora caso haja demanda para tal mudança (veja por exemplo Gordon, 1989; Gordon; Rosengren; Sundstrom, 1992), inclusive por intermédio da utilização de semioquímicos conforme evidência experimental (Greene; Gordon, 2003).

Portanto, tal como no caso dos seres humanos onde nossa posição na escala social não é função linear de uma simples determinação

genética, embora possa ser influenciada por nossa individualidade genética (o que pode ir desde a sermos filhos biológicos de pessoas cuja posição social é transmitida hereditariamente – como na realeza de muitos países – , até a nossa probabilidade de ocupar uma posição social ou outra em contextos de mobilidade social ser afetada por nossos traços cognitivos e comportamentais, sendo que as diferenças nesses traços entre as pessoas são parcialmente explicáveis pela diferença genética entre elas, conforme achados da genética comportamental⁹), o mesmo pode ser dito dos insetos eussociais, ao menos por referência ao caso das formigas.

A diversidade existente entre diferentes espécies de formigas eussociais a respeito das características de suas trilhas químicas (por exemplo, em algumas tendo um caráter mais fixo, outras tendo um caráter mais transiente) demonstra que diferentes soluções foram favorecidas pela seleção natural a depender do ambiente.

Nossos traços cognitivos e comportamentais continuam sujeitos à seleção natural (embora em sociedades industriais modernas isso seja menos evidente que em sociedades de pequena escala com fertilidade natural). Contudo, a vasta diferença estrutural entre nós e os insetos eussociais faz com que o impacto dessa seleção natural seja bem diferente entre nós e eles, uma vez que a mudança intergeracional dos insetos eussociais é mais rápida, seus sistemas nervosos mais simples e sua existência evolutiva é muito mais longa que a do gênero *Homo* (uma diferença entre aproximadamente 150 milhões de anos e 2 milhões de anos, respectivamente). Portanto, uma comparação entre a nossa sociabilidade e a dos insetos eussociais precisa levar em conta essas diferenças, não para negar o caráter social genuíno das manifestações

⁹ Vide Plomin et al (2010).

desses insetos, mas sim para entender como o caráter social genuíno pode se manifestar de maneiras diferentes que na espécie humana e ainda assim remanesça possível traçar um paralelo entre organismos vivos tão diferentes.¹⁰

Um viés que afeta essa questão diz respeito ao fato de o *Homo sapiens* ter sido o único hominídeo a persistir entre a miríade de outros hominídeos que fizeram parte da cadeia evolutiva desse ramo de primatas nos últimos milhões de anos. Caso ainda existissem hominídeos comparáveis ao *Homo erectus* por aí, tendo uma aparência morfológica e funcionamento fisiológico muito similares aos do *Homo sapiens*, mas com um funcionamento cognitivo que presumivelmente veríamos como mais similar a de grandes primatas antropóides não-humanos ao invés do nosso, teríamos mais evidente diante de nós o quanto da nossa organização social é ela mesma determinada geneticamente (mas com flexibilidade como já destacado). Porque se pudéssemos nos comparar com seres vivos de morfologia e fisiologia similares a nossa, mas com outro tipo de comportamento (inclusive social) – distinto e ainda assim com várias similaridades ao nosso, embora sem nosso mesmo funcionamento cognitivo –, seria mais evidente que a diferença de organizações sociais entre nós e eles se deveria à evolução por seleção natural sob a base genética subjacente das respectivas espécies. Isso é válido mesmo assumindo que haja uma variedade de organizações sociais entre os *Homo sapiens* (como de fato há) e uma variedade de organizações sociais nessa hipotética espécie de

¹⁰ Compare como podemos traçar um paralelo na evolução de asas em insetos e em morcegos, onde as asas não são fruto de descendência comum – homologia –, mas sim de convergência funcional – analogia – entre seres vivos distantes filogeneticamente. Lorentz (1981) comenta que um dos princípios originais subjacentes à fundação da Etologia como disciplina biológica seria o estudo do comportamento em termos de homologia e analogia, tal como já era feito para órgãos e tecidos corporais.

hominídeo não-sapiens que poderia ainda existir, já que a questão pode ser entendida em termos da diferença entre como variam as organizações sociais dessas espécies hominídeas dado os ambientes de referência. Com tantas espécies de insetos eussociais, mesmo quando comparadas apenas formigas entre si ou apenas abelhas entre si e assim por diante, obviamente é muito mais fácil proferirmos esse juízo a respeito delas.

Diante dos pontos acima apresentados, o objetor poderia procurar algum argumento biológico – assumindo agora que tanto seres humanos como insetos eussociais estão sujeitos aos mesmos princípios norteadores da evolução da vida – que pudesse refinar sua diferenciação entre esses casos sob bases não-arbitrárias.

A seguinte objeção mais cientificamente refinada poderia ser apresentada. A eussocialidade dos insetos eussociais se originou e persiste na evolução por seleção natural em termos de haplodiploidia¹¹ (válido para ordem *Hymenoptera*, não se aplica aos isópteros/cupins¹²), sendo caso clássico de seleção de parentesco, onde o comportamento altruísta evolui em razão do benefício reprodutivo para indivíduos relacionados geneticamente ao agente, na medida de sua relação genética, conforme dado na Regra de Hamilton ($R > c/b$): se o benefício reprodutivo que o comportamento de um indivíduo confere para outro

¹¹ Mesmo sendo reconhecido hoje que a haplodiploidia tenha desempenhado apenas um papel menor na evolução da eussocialidade (Abbot et al, 2011), uma vez que há animais eussociais diploides e a haplodiploidia não tem uma conexão estatisticamente forte com a eussocialidade (Nowak; Tarnita; Wilson, 2010, com base em Wilson, 2008; Linksvayer; Wade, 2005; Queller; Strassmann, 1998), ainda se segue que a haplodiploidia é uma característica comum nos insetos eussociais da ordem *Hymenoptera* (nosso foco aqui) não compartilhada pelos primatas humanos. Dessa forma, a argumentação feita a seguir em termos de ser uma característica dissimilar entre insetos eussociais e primatas humanos ainda se segue mesmo admitido papel explanatório reduzido para esse fator.

¹² Os insetos eussociais da ordem *Hymenoptera*, que engloba formigas, abelhas e vespas, são hadiplóides, enquanto os insetos eussociais da subordem *Isoptera* (ordem *Blattodea*), que engloba os cupins, são diplóides tal como nós.

multiplicado pelo coeficiente de relação genética (*relatedness*) entre eles supera o custo reprodutivo para o indivíduo que realiza referido comportamento, então esse comportamento será selecionado em termos de *fitness* inclusivo (sucesso reprodutivo inclusivo). As operárias em muitas espécies eussociais (mas não em todas) compartilham entre si mais que metade de sua base genética, sob coeficiente de 0,75. Assim, as operárias ainda maximizariam seu sucesso reprodutivo de forma indireta, mesmo abdicando da reprodução.¹³

Diante desse peso da seleção de parentesco e da haplodiploidia de insetos eussociais (especialmente da ordem *Hymenoptera* como já explicado), o objetor hipotético poderia argumentar que isso é uma dissimilaridade muito evidente para com o caso dos primatas humanos, cuja evolução por seleção natural se dá em termos de diploidia, e envolve cooperação entre estranhos não-geneticamente relacionados. Na literatura de evolução da cooperação inclusive se destaca o fato de que, entre os animais, a reciprocidade é rara ou inexistente, enquanto no ser humano a reciprocidade é componente-chave de nossa sociabilidade. Note que existem outros animais não-humanos que apresentam cooperação entre não-parentes, mas a evolução da cooperação com base em benefício mútuo entre não-parentes nos animais não-humanos tende a ser produzida por outros mecanismos que não a reciprocidade (West; El Mouden; Gardner, 2011). No caso em apreço, dos insetos eussociais, entretanto, a questão da relação genética entre os membros da colônia seria muito acentuada, com a seleção de

¹³ Também são necessárias algumas qualificações aqui. Há diversas espécies formigas eussociais nas quais as operárias retém a capacidade de pôr ovos. Especificamente em espécies nas quais a relação genética entre as operárias é menor (quando a rainha acasala com mais de um macho, reduzindo o coeficiente para 0,25), a teoria permite prever e os dados corroboram que há policiamento por parte das demais operárias para impedir essa reprodução “desviante” (Ratnieks 1988; Wensellers; Ratnieks 2006). A harmonia reprodutiva surgiria, contraintuitivamente, de uma relação genética menor, mas ainda dentro do escopo preditivo da teoria do *fitness* inclusivo.

parentesco sendo o carro-chefe de sua organização social colonial, o que só aumentaria sua distância em relação ao caso dos primatas humanos, marcado pela cooperação de não-parentes com base em reciprocidade (para além da seleção de parentesco que, não obstante, também existe no caso humano). Assim, concluiria nosso objetor, poderíamos negar o status de entidade social às trilhas químicas das formigas ou abelhas eussociais, ou negar que haja ontologia social entre os insetos eussociais, em razão de seu comportamento aparentemente social não ter as mesmas bases genéticas e evolutivas que o do comportamento social genuíno, o nosso.

Contudo, entendo que essa objeção também falha. Primeiro, se poderíamos admitir que o comportamento genuinamente social pode variar em termos de suas causas próximas (por exemplo, instanciado por organismos com sistemas neuronais bem distintos do nosso, talvez mesmo por uma química baseada em silício ao invés de carbono), também parece que devemos admiti-la no que diz respeito às suas causas remotas em termos da evolução da cooperação.

Comportamento genuinamente social vem a existir se vier a existir cooperação entre indivíduos de uma mesma espécie sob uma vida grupal com base na evolução por seleção natural (ou outro processo seletivo).¹⁴ Mas como especificamente foi a base de seleção para essa

¹⁴ A condição de seleção aqui diz respeito ao organismo ter o design adaptativo para a cooperação, de tal modo que seu sucesso reprodutivo esteja parcialmente entrelaçado ao sucesso reprodutivo de outros organismos da mesma espécie. Isso afasta o problema de um comportamento coordenado de forma meramente accidental. Em publicação sob análise dos pareceristas, defendo que a função biológica em termos de efeitos seletivos consegue satisfazer o mesmo papel teórico que a intencionalidade mental na ontologia social padrão, no sentido de distinguir comportamento cooperativo genuinamente social subjacente à ontologia social do comportamento coordenado de forma meramente accidental. Parte desse argumento já encontrava-se esboçado na minha tese de doutorado (Brito Jr., 2021). A função em termos de efeito seletivo como a preconizo também admite outros processos seletivos que não a seleção natural, mas entendo que a seleção natural ocupa um papel de destaque nesse âmbito. Entretanto, para fins do presente artigo, se alguém discordar de minha posição por relativizar o papel da seleção natural ou outros processos seletivos aqui (p. e. argumentando que bastaria uma condição mais fraca de função biológica em termos de papel causal que independe

cooperação vir a existir? A literatura da evolução da cooperação nos confere essa resposta, mostrando diferentes caminhos pelos quais isso se desenvolve evolutivamente, bem como a partir dela pode-se definir a cooperação em termos de comportamentos que evoluíram em razão de seu efeito sobre o sucesso reprodutivo de indivíduos outros que não aquele que realiza o comportamento (Birch, 2017, p. 34).¹⁵ Não haveria uma motivação independente para restringir a sociabilidade genuína apenas para o caso da evolução da cooperação em mútuo benefício via reciprocidade, e excluir a seleção de parentesco.

Além disso, como demonstrado por Birch (2017, p. 55-58), a evolução da cooperação em mútuo benefício ainda segue a Regra de Hamilton como um princípio organizador geral, considerados casos de explicações de *fitness* direto onde *c* torna-se negativo, via reciprocidade, punição, entre outros ali discutidos. Portanto, há razões adicionais para entendermos que os mesmos princípios subjacentes operam tanto no caso dos insetos

da causação remota por inteiro), não haverá nenhum problema para o mérito deste argumento, ao contrário, porque nesse caso o interlocutor estaria adotando uma posição ainda mais plural em relação à causação e constituição da cooperação animal (e de outros organismos vivos). Sobre a distinção entre função biológica de efeitos seletivos e de papel causal, veja Garson, 2016.

¹⁵ Na definição técnica completa proposta por Birch: "Cooperação biológica: uma ação A, realizada como parte de uma estratégia S em uma tarefa T, é cooperativa se e somente se S tiver sido, em sua história recente, mantido pela seleção em parte por conta do efeito positivo de realizar A em T sobre o sucesso reprodutivo de outros organismos." Ele também considera que devemos distinguir entre essa noção técnica de cooperação e uma noção mais informal de cooperação: "*We should take care to distinguish the theoretical notion of cooperation, which I define in terms of recent selection history, from the more intuitive idea of an action that benefits other organisms. Obviously, the discovery that natural selection can explain the maintenance of actions that benefit other organisms, even at a cost to the actor, was a non-trivial breakthrough, in this case one properly credited to Hamilton (1964). But, as with adaptation, this discovery now provides us with a rationale for introducing a theoretical notion of cooperation in which the role of selection is made explicit. This theoretical notion is an explanatory category: to classify a behaviour as cooperative is to assign it to a class of behaviours whose maintenance under selection is explained by their positive effect on the reproduction of other organisms. It is also helpful to stress again here the distinction between a trait's origin and its recent maintenance. It may be a definitional truth that a behaviour that is cooperative has been maintained by selection in recent history, but it remains, for any particular cooperative behaviour, a substantive explanatory hypothesis that selection explains the origin of that behaviour. It therefore remains a substantive explanatory hypothesis that the behaviour is an adaptation (rather than, say, an 'exaptation' in the sense of Gould and Vrba, 1982).*" (Birch, 2017, p. 24)

eussociais como dos primatas humanos, mesmo que sob diferentes trajetórias evolutivas, de forma que em todos esses casos trata-se da evolução de sociabilidade genuína. Em decorrência disso, tanto a seleção por parentesco atua em humanos, como a evolução da cooperação em benefício mútuo também atua em insetos eussociais, apenas que não por reciprocidade, como no caso das vespas primitivamente eussociais do gênero *Polistes*, entre as quais as rainhas são regularmente agressivas em relação às operárias inativas, e a remoção experimental das rainhas diminui a atividade das operárias (Clutton-Brock; Parker, 1995, p. 14, citado por Birch, 2017, p. 57-58). De fato, o framework hamiltoniano, como incorporado na ecologia comportamental (quando o comportamento assim estudado é o humano, falamos em Ecologia Comportamental Humana, praticada por bioantropólogos; Nettle et al, 2013; Winterhalder; Smith, 2000), é um dos paradigmas mais proeminentes internacionalmente para estudos de caçadores-coletores em antropologia (Bettinger; Garvey; Tushingham, 2015; Kelly, 2013).¹⁶

Segundo, mesmo para os insetos eussociais, há casos em que há aparente cooperação entre indivíduos não-aparentados em grupos enormes, quando se diz respeito ao fenômeno da supercolonialidade, que pode alcançar escalas monumentais, como no caso da supercolônia japonesa de *Formica yessensis* estimada em 306 milhões de operárias para 1 milhão de rainhas dispersas ao longo de 2,5km² entre 45.000 ninhos interconectados (Higashi; Yamauchi, 1979; Gullan; Cranston, 2005, p. 300).

¹⁶ Aqui destaco os caçadores-coletores porque o registro etnográfico a seu respeito perfaz o campo de teste mais fundamental para a teoria evolutiva no comportamento humano, visto serem sociedades de pequena escala com fertilidade natural, e, nesse sentido, fornecendo a evidência (não-paleoantropológica) mais próxima do comportamento ancestral do *Homo sapiens* anteriormente ao advento da agricultura. Contudo, afirmo isso sem prejuízo de que não podemos simplesmente equivaler os modos de vida dos caçadores-coletores recentes ou atuais – já afetados pelo contato com sociedades agrícolas, pastoris e industriais de larga escala – com os caçadores-coletores como existiram previamente à invenção da agricultura ou ao contato com sociedades de maior escala, algo que é amplamente reconhecido na ecologia comportamental humana (Marlowe, 2005).

As supercolônias tem como características distintivas a polidomia (com múltiplos ninhos socialmente conectados, mas fisicamente separados) e a poliginia (com múltiplas rainhas reprodutivas), sendo que as operárias circulam livremente entre os ninhos. Isso é marcante na medida em que o coeficiente de relação genética cai para zero quando considerado o conjunto da obra, bem como pela relação teórica forte entre monogamia e eussocialidade (havendo múltiplas evidências de que a monogamia seria ao menos o estado ancestral da eussocialidade; Hughes et al, 2008). Vale a pena conferir o esquema em Helanterä (2022) sintetizando em quais famílias de formigas há evidências do fenômeno da supercolonialidade, bem como sua discussão a respeito da definição de supercolonialidade e unicolonialidade.

Enquanto talvez não se trate de cooperação no sentido técnico da evolução da cooperação (no caso de a ocorrência da supercolonialidade ser subproduto ao invés de efeito seletivo), ainda se trata de cooperação no sentido mais amplo de que as colônias agrupadas em supercolônias saem favorecidas por compartilharem espaços e operárias, ao invés de se confrontarem ou se evitarem mutuamente, bem como pelo fato de envolverem seres vivos que em si mesmos já são eussociais, seu comportamento cooperativo eusocial sendo então realizado num contexto supercolonial. De um lado, Helanterä (2022; cf. também Helanterä et al., 2009; Helanterä, 2009) discute como a seleção de parentesco ainda pode ser aplicada em escalas espaciais menores no contexto supercolonial: a falta de fronteiras territoriais e a mistura de indivíduos não significa que há ausência de estrutura de parentesco, por exemplo, isolamento por distância devido à viscosidade da população pode significar que os indivíduos nos ninhos vizinhos sejam geneticamente mais similar uns aos outros do que indivíduos nas pontas em extremos opostos da supercolônia. Por outro, também pode ser útil

recorrer à construção de nicho social para entender esse fenômeno, na medida em que traços sociais coevoluem e modificam o regime seletivo uns dos outros, de modo que a estrutura de parentesco que fornece o regime seletivo para os traços sociais pode também ela mesma depender de outros traços sociais, como decisões de dispersão e acasalamento (Helanterä, 2022; Ryan et al, 2016). Modelos de construção de nicho têm sido propostos para o entendimento de como a cultura afeta a evolução humana (Laland; Odling-Smee; Feldman, 2001; Laland; Matthews; Feldman, 2016). Sob modelos de construção de nicho, o ser vivo consegue modificar parcialmente suas pressões adaptativas por modificar seu ambiente, e, no caso de uma construção de nicho social, essa modificação do ambiente opera por intermédio de variáveis sociais (*lato sensu*).

De forma interessante para a questão anteriormente discutida sobre determinações genéticas da organização social, dentro de uma mesma espécie, como *Formica exsecta*, pode-se encontrar tanto populações supercoloniais como populações coloniais no sentido padrão. De fato, em muitas espécies, populações monogâmicas e poligâmicas podem mesmo coexistir no mesmo habitat. Então, nessas espécies há variação em organização social entre diferentes colônias.

O objeto poderia então levantar a questão de algumas divergências recentes no ramo da biologia evolutiva teórica. Edward O. Wilson junto a Martin Nowak e Corina Tarnita (2010) argumentaram contra o framework da seleção de parentesco e da Regra de Hamilton como produzindo apenas um trabalho matemático abstrato sem firme base na observação de campo e na genômica. Na questão da evolução da eussocialidade, eles alegam que uma visão padrão de seleção natural (ao invés do modelo do *fitness* inclusivo) mais algumas considerações sobre estrutura populacional seriam suficientes para dar conta do recado. Além disso, eles buscaram traçar um passo a passo do processo causal

de como se chegaria à eussocialidade tomando por referência os insetos eussociais. A crítica deles ao modelo do *fitness* inclusivo levou a uma resposta por 137 cientistas teóricos e experimentais do campo da biologia evolutiva (Abbot et al, 2011), na qual estes reafirmaram o valor preditivo da teoria do *fitness* inclusivo entre vários campos de pesquisa biológica, observacional e experimental.

Enquanto eu endosso o modelo do *fitness* inclusivo, vale a pena explorar o que aconteceria se adotássemos o modelo proposto por Wilson, Nowak e Tarnita. O objetor teria, então, um terreno melhor para inserir sua crítica ao que estou propondo aqui? Não vejo sob melhor termo. Sob esta visão alternativa, o papel da relação genética se torna menos crucial para uma distinção humano *versus* insetos eussociais, uma vez que a relação genética passa a ser vista como consequência ao invés de causa da eussocialidade. Também afirmam em seu artigo que os diferentes papéis dos membros reprodutivos e dos membros não-reprodutivos das colônias eussociais não seriam geneticamente determinados/fixados, sendo produtos dos mesmos genes ou de conjuntos de genes com flexibilidade fenotípica (fenótipos alternativos do mesmo genótipo, no que diz respeito à casta e divisão do trabalho, conquanto possam diferir livremente em outros genes), o que só reforçaria meu ponto anteriormente discutido contra visões simplistas da determinação genética subjacente ao comportamento. Por fim, eles também abordam a ideia de tratar a colônia eusocial como um superorganismo, ideia também explorada por biólogos que aceitam o modelo do *fitness* inclusivo (como Helanterä, 2016). Essa última questão precisa de um tratamento um pouco mais detido.

Sob a visão de Wilson, Nowak e Tarnita (2010), e como também explicado por Wilson em subsequente livro de divulgação científica (2012), em certas espécies eussociais mais extremas (presumivelmente:

eussocialidade obrigatória, altamente especializada, especialmente quando as operárias perdem os ovários^{17 18}), a colônia pode ser entendida como um Superorganismo: nesse caso, a seleção natural opera sobre um indivíduo genético correspondente à própria colônia, qual seja, o genoma da rainha, sendo as operárias uma extensão fenotípica extrassomática do genoma da rainha. No linguajar metafórico que Wilson adota no livro, seriam como extensões robóticas da rainha. Então, tal como as células de um organismo multicelular se diferenciam em reprodutivas e não-reprodutivas, ativando ou desativando determinadas características para que possam dar forma a tecidos diferentes, e assim por diante, a partir de um mesmo programa genético, as operárias aqui poderiam ser vistas de maneira similar, como satisfazendo um papel dirigido ao todo. Com isso, as operárias não cooperariam entre si para maximizar seu *fitness* inclusivo, mas servem como partes operando na maximização do *fitness* do superorganismo colonial, cujo genótipo corresponde ao da rainha, portanto, maximização de *fitness* individual.

Aqui o objeto poderia alegar que no caso da eussocialidade plena não se trata mais de cooperação e, logo, estamos fora do âmbito da ontologia social. Não se negaria ontologia social a todos os insetos eussociais, mas àqueles que alcançaram o estágio final que leva à

¹⁷ Os autores aduzidos deixam um tanto ambíguo sob qual nível de eussocialidade se chega ao superorganismo, apenas ficando claro que não enquadrariam espécies primitivamente eussociais nisso. Por isso sugeri este cenário "extremo" ou "pleno" como aquele em que mais faria sentido ver as operárias como extensões extrassomáticas da rainha. Talvez os autores adotem um critério mais relaxado, em razão de não trabalharem no framework do *fitness* inclusivo (e assim evitam abordar o conflito reprodutivo entre operárias e rainhas), contudo, ao menos para fins do propósito do presente artigo, entendo que essa restrição seria mais plausível.

¹⁸ Na maioria das espécies de formigas, as operárias retêm os ovários, mas sofrem restrições reprodutivas, relativos a comportamento ou desenvolvimento. O caso mais extremo de restrição reprodutiva via desenvolvimento envolve a perda completa dos órgãos reprodutivos pelas operárias (Khila; Abouheif, 2010).

transição para o superorganismo. O problema que vejo com esta objeção é que, como a colônia ainda seria composta por seres vivos individualmente autônomos capazes de ação de forma similar ao das colônias que não sejam plenamente eussociais (isto é, usando os mesmos mecanismos ou muito próximos), então mesmo sob uma visão de superorganismo nestes termos ainda entenderíamos a atividade de cada indivíduo da colônia como suficientemente distinta (por exemplo, tendo uma *guidance* autônoma ao nível de cada organismo, como na abordagem de ação primitiva de Jones, 2016) e, assim, apta à atribuição de ontologia social. A metáfora com robôs feita por Wilson (2012) é enganadora, porque – mesmo se não fora sua intenção – nos leva a pensar em entidades sem propósito biológico autônomo, que simulam comportamentos porque foram programados para fazê-lo. As questões levantadas pela robótica como construção intencional de artefatos capazes de comportamento ou de simulação de comportamento são de outra ordem. Conquanto eu não esteja formulando aqui um argumento mais decisivo – a questão do superorganismo e da ontologia social mereceria tratamento em artigo separado –, me parece que *prima facie* a objeção não é contundente.

Considerações finais

Tendo discutido como os princípios pertinentes de biologia teórica, conjugados à evidência empírica pertinente, favorecem que tratemos em par os casos de insetos eussociais e primatas humanos, ao objetor restaria uma última tentativa: afirmar que apenas seres vivos que tenham estados mentais intencionais de cognição superior como os nossos poderiam ter ontologia social¹⁹. Mesmo que o objetor não

¹⁹ Note que a tese da dependência da mente em ontologia social diz respeito a estados mentais intencionais, não à questão da experiência fenomênica/qualitativa (e nem à noções sensoriomotoras). Consulte o 2º capítulo da minha tese de doutorado para uma discussão desse ponto. Nesse sentido,

exigisse intencionalidade coletiva como Searle o faria, ainda exigiria a aptidão para ter estados mentais intencionais de maneira geral. Voltaríamos à questão da tese da dependência da mente em ontologia social, que discuti à exaustão em minha tese de doutorado (Brito Jr., 2021). A resposta definitiva dependeria de avanços na psicologia animal (para uma revisão panorâmica das questões teóricas envolvidas nessa empreitada empírica, veja Andrews, 2014).

Contudo, considero insatisfatória tal atitude teórica, porque: 1) considero implausível que uma noção de atitude proposicional/estado mental intencional feita sob medida para o caso dos primatas humanos possa se aplicar aos insetos eussociais sem trivializar o que se entenderia por essa intencionalidade mental e ainda assim estivesse apta a satisfazer a motivação subjacente à tese de que tipos sociais dependeriam de estados mentais intencionais (a meu ver, a motivação residiria na capacidade de ao menos potencialmente gerar representações dos tipos sociais, mesmo que não tenhamos representações cognitivas de todos os tipos sociais prontamente disponíveis²⁰); 2) entendo que é mais correto afirmarmos que já sabemos que certos animais são sociais, no caso em apreço, que insetos eussociais são animais sociais, não sendo necessário esperar avanços a respeito da questão da intencionalidade mental para conseguirmos afirmar isso.

Falando em termos gerais, para que haja ontologia social o que se faz necessário é que haja certo tipo de comportamento, sujeito aos princípios relativos à evolução da cooperação, mesmo na ausência de

mesmo a questão de se colônias de formigas ou outros insetos eussociais podem ser ditas ou não como possuindo experiência fenomênica coletiva não interfere a princípio na questão da tese da dependência da mente como discutida em ontologia social.

²⁰ Consulte aqui a discussão feita por Thomasson (2003) e Khalidi (2015), ou o 1º e 2º capítulo de minha tese de doutorado, nos quais busquei descrever com maior precisão o que seria a tese da dependência da mente em ontologia social.

uma intencionalidade mental de cognição superior como a nossa. Sob esse critério, os insetos eussociais claramente passam. E me parece que, depois de tudo que revisei anteriormente a respeito do paralelo entre o caso dos primatas humanos e dos insetos eussociais, insistir que insetos eussociais só poderiam ter ontologia social se tiverem intencionalidade mental em sentido comparável ao nosso seria fechar os olhos para a forma como o fenômeno social é explicável por meio de princípios relativos a todos os seres vivos do planeta, não apenas àqueles aptos à cognição superior. Assim, nossa arregimentação metafísica da realidade como um todo faria bem em reconhecer a ontologia social de insetos eussociais, com entidades sociais conforme cada espécie e colônia.

Referências

- ABBOT, P.; ABE, J; ALCOCK, J. *et al.* Inclusive fitness theory and eusociality. *Nature* 471(7339): 2011, E1-4; author reply E9-10.
- AMDAM GV, CSONDES A, FONDRK MK, PAGE RE. Complex social behaviour derived from maternal reproductive traits. *Nature* 439, 2006, p. 76-78.
- ANDREWS, Kristin. *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. New York: Routledge, 2014.
- BATRA, Suzanne W. T. Nests and Social Behavior of Halictine bees of India (Hymenoptera: Halictidae). *The Indian Journal of Entomology* 28 (3): 1966, p. 375-393.
- BETTINGER, Robert L.; GARVEY, Raven; TUSHINGHAM, Shannon. *Hunter-Gatherers: Archaeological and Evolutionary Theory*. New York: Springer, 2015.
- BIRCH, Jonathan. *The Philosophy of Social Evolution*. Oxford: OUP Oxford, 2017.
- BRITO JÚNIOR, Valdenor M. *Ontologia Social Não-Mentalista: Uma Abordagem Comportamental Biocentrada*. (Tese de Doutorado) Florianópolis: UFSC, 2021.
- CAMPBELL, D.T. The two distinct routes beyond kin selection to ultrasociality: Implications for the Humanities and Social Sciences. In: BRIDGEMAN, D. (ed.) *The*

- Nature of Prosocial Development: Theories and Strategies*. New York: Academic Press, 1983, p. 11-39.
- CHIURLIZA, B.; ROGERS, M.; SCHNEIDER, M.; CHU, C.; JOINER, T. Evolutionary Processes in Suicide. *Curr Opin Psychol*. 22: 2018, p. 84–88.
- CLUTTON-BROCK TH, PARKER GA. Punishment in animal societies. *Nature* 373(6511); 1995, p. 209-16.
- FOSTER, KR Foster; RATNIEKS, FLW. A new eusocial vertebrate? *Trends Ecology Evolution* 20, 2005, p. 363–364.
- GARG, Ketika; PADILLA-IGLESIAS, Cecilia; RESTREPO, Ochoa Nicolás; KNIGHT, V. Bleu. Hunter–gatherer foraging networks promote information transmission. *Soc. Open Sci.* 8.211324, 2021.
- GARSON, Justin. *A Critical Overview of Biological Functions*. New York: Springer, 2016.
- GEDEON, Péter. Spontaneous Order and Social Norms Hayek's Theory of Socio-Cultural Evolution. *Society and Economy* Vol. 37, No. 1, 2015, p. 1–29.
- GORDON DM. Dynamics of task switching in harvester ants. *Anim Behav*. 38: 1989, p. 194–204.
- GORDON, Deborah. *The Ecology of Collective Behavior*. *PLOS Biology* 12(3), 2014.
- GORDON DM, ROSENGREN R, SUNDSTROM L. The allocation of foragers in red wood ants. *Ecol Entomol* 17: 1992, 114–120.
- GORDON DM, GUETZ A, GREENE MJ, HOLMES S. Colony variation in the collective regulation of foraging by harvester ants. *Behav Ecol*. 22(2): 2011, p. 429-435.
- GORDON, Deborah. *The Ecology of Collective Behavior*. Princeton: Princeton University Press. 2023.
- GREENE MJ, GORDON DM. Cuticular hydrocarbons inform task decisions. *Nature*. 2003, p. 423-432.
- GULLAN, P. J.; CRANSTON, P. S. *The Insects. An Outline of Entomology*, third ed, P.J. Oxford: Blackwell, 2005.

HAMILTON WD. The genetical evolution of social behaviour, I, II. *J Theor Biol.* 7: 1964, p. 1–52.

HAYEK, F. A. *Studies in Philosophy, Politics and Economics*. London: Routledge & Kegan Paul, 1967.

HELANTERÄ, H. Do unicolonial wood ants favor kin? *Journal of Biology* 8: art. 56, 2009.

HELANTERÄ, H., STRASSMANN, J.E., CARRILLO, J. & QUELLER, D.C. Unicolonial ants: where do they come from, what are they and where are they going? *Trends in Ecology & Evolution* 24: 2009, p. 341–349.

HELANTERÄ, H. An organismal perspective on the evolution of insect societies. *Frontiers in Ecology and Evolution* 4: art. 6, 2016.

HELANTERÄ, H. Supercolonies of ants (Hymenoptera: Formicidae): ecological patterns, behavioural processes and their implication for social evolution. *Myrmecological News* 32: 2022, p. 1–22.

HUGHES, William O. H.; OLDROYD, Benjamin P.; BEEKMAN, Madeleine; RATNIEKS, Francis L. W. Ancestral Monogamy Shows Kin Selection Is Key to the Evolution of Eusociality. *Science* 320 (5880): 2008, p. 1213–1216.

HIGASHI, Seigo; YAMAUCHI, Katsusuke. Influence of a Supercolonial Ant *Formica* (*Formica*) *Yessensis* Forel on the Distribution of Other Ants in Ishikari Coast: Biological and ecological studies of a supercolonial ant *Formica yessensis* Forel.X. *Japanese Journal of Ecology* 29(3), 1979, p. 257–264.

JANG H, BOESCH C, MUNDRY R, BAN SD, JANMAAT KRL. Travel linearity and speed of human foragers and chimpanzees during their daily search for food in tropical rainforests. *Sci. Rep.* 9, 2019, p. 1–3.

JOINER TE, HOM MA, HAGAN CR, SILVA C. Suicide as a derangement of the self-sacrificial aspect of eusociality. *Psychological Review* 123: 2016, p. 235–254.

JONES, Derek. *The Biological Foundations of Action*. New York: Routledge, 2016.

KELLY, Robert L. *The Lifeways of Hunter-Gatherers: The Foraging Spectrum*. Cambridge University Press, 2013.

- KHALIDI, Muhammad Ali. Three Kinds of Social Kinds. *Philosophy and Phenomenological Research* 90, 2015, p. 96-112.
- KHILA A, ABOUHEIF E. Evaluating the role of reproductive constraints in ant social evolution. *Philos Trans R Soc Lond B Biol Sci.* 365(1540): 2010, p. 617-30.
- LALAND KN, ODLING-SMEE J, FELDMAN MW. Cultural niche construction and human evolution. *J Evol Biol.* 14(1): 2001, p. 22-33.
- LALAND K, MATTHEWS B, FELDMAN MW. An introduction to niche construction theory. *Evol Ecol.* 30: 2016, p. 191-202.
- LAYTON R, O'HARA S. Human social evolution: a comparison of hunter-gatherer and chimpanzee social organization. *Proc. Br. Acad.* **158**, 2010, p. 83-113.
- LORENTZ, Konrad. *The Foundations of Ethology*. New York: Springer Verlag, 1981.
- LINKSVAYER TA, WADE MJ. The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. *Q Rev Biol.* 80: 2005, p. 317-336.
- NETTLE, Daniel; GIBSON, Mhairi A.; LAWSON, David W.; SEA, Rebecca. Human behavioral ecology: current research and future prospects. *Behavioral Ecology* 24:5, 2013, p. 1031-1040.
- NOWAK MA, TARNITA CE, WILSON EO. The evolution of eusociality. *Nature.* 466: 2010, p. 1057-1062.
- PEETERS, C.; R. CREWE. Insemination controls the reproductive division of labour in a ponerine ant. *Naturwissenschaften* 71: 1984, p. 50-51.
- PEETERS, C. The occurrence of sexual reproduction among ant workers. *Biological Journal of the Linnean Society* 44: 1991, p. 141- 152.
- PEIXOTO, A. V., CAMPIOLO, S., LEMES, T. N., DELABIE, J. H. C., HORA, R. R.. Comportamento e estrutura reprodutiva da formiga *Dinoponera lucida* Emery (Hymenoptera, Formicidae). *Revista Brasileira De Entomologia* 52(1), 2008, p. 88-94.

PLOMIN, Robert; DEFRIES, John C.; MCCLEARN, Gerald E.; MCGUFFIN, Peter. *Genética do Comportamento*. Porto Alegre: Artmed, 2010.

QUELLER DC, STRASSMANN JE. Kin selection and social insects. *Bioscience* 48: 1998, p. 165–175.

RATNIEKS, F.L.W. Reproductive harmony via mutual policing by workers in eusocial Hymenoptera. *American Naturalist* 132: 1988, p. 217–236.

ROBINSON GE, FERNALD RD, CLAYTON DF. Genes and social behavior. *Science* 322: 2008, p. 896–900.

RYAN, P.A.; POWERS, S.T.; WATSON, R.A. Social niche construction and evolutionary transitions in individuality. *Biology & Philosophy* 31: 2016, p. 59–79.

SEARLE, John. *The Construction of Social Reality*. New York: Free Press, 1995.

SEARLE, John. *Making the Social World: The Structure of Human Civilization*. Oxford: Oxford University Press, 2010.

THOMASSON, Amie. Foundations for a Social Ontology. *Protosociology* 18, 2003, p. 269–290.

TURCHIN, Peter. The Puzzle of Human Ultrasociality: How Did Large-Scale Complex Societies Evolve? In: RICHERSON, P. J.; CHRISTIANSEN, M. H. (eds) *Cultural Evolution*. Cambridge: MIT Press, 2013, p. 67–73.

WENSELEERS, T.; RATNIEKS, F.L.W. Comparative analysis of worker reproduction and policing in eusocial Hymenoptera supports relatedness theory. *The American Naturalist* 168: 2006, E163–E179.

WEST, S. A.; EL MOUDEN, C.; GARDNER, A. Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior* 32:4, 2011, p. 231–262.

WILSON, Edward O. *Sociobiology: The New Synthesis*. Harvard: Harvard University Press, 1975.

- WILSON, Edward O.; HÖLLDOBLER, Bert. Eusociality: Origin and consequences. *PNAS* 102:38, 2005, p. 13367-13371.
- WILSON, Edward O. *The Social Conquest of Earth*. New York: Liveright Publishing/W. W. Norton and Company, 2012.
- WILSON, E. O.; NOWAK, M. A. Natural selection drives the evolution of ant life cycles. *Proceedings of the National Academy of Sciences of the United States of America* 111(35): 2014, p. 12585-12590.
- WINTERHALDER, Bruce; SMITH, Eric Alden. Analyzing Adaptive Strategies: Human Behavioral Ecology at Twenty-Five. *Evolutionary Anthropology* 9:2, 2000, p. 51-72.
- YAN H, LIEBIG J. Genetic basis of chemical communication in eusocial insects. *Genes Dev.* 35(7-8): 2021, p. 470-482.

7

COMO SER REALISTA SOBRE MENTE DE GRUPOS?

Luiz Paulo Da Cas Cichoski

Entidades sociais parecem compor a realidade. É difícil negar que fazemos parte de Estados Nacionais, que pagamos impostos (com dinheiro), que as pessoas se casam, jogam xadrez e que times de futebol fazem gols e vencem partidas. Algumas porções dessa realidade social são unidades coletivas, ou seja, entidades constituídas por um conjunto de indivíduos. Instituições, organizações e grupos parecem figurar como agentes da realidade social. Empresas contratam funcionários, Universidades emitem diplomas, Estados Nacionais entram em guerras, Supremas Cortes julgam a constitucionalidade de leis. Será que essas caracterizações sobre a realidade social são precisas? Será que nosso discurso sobre entidades sociais são literais? Ou esse discurso é só um “modo de falar”, isto é, essas sentenças são metafóricas?

No presente texto, vou discutir um pressuposto central para a legitimidade de sentenças onde entidades sociais figuram como sujeitos. Muitas sentenças só podem ser verdadeiras se os seus sujeitos possuem estados mentais: “Maria lançou a bola intencionalmente” apenas se tinha a intenção de fazer isso; “João sabe que o Brasil está no Hemisfério Sul” apenas se acredita nisso. Apesar de menos comum do que atribuições de ações como as do parágrafo anterior, também fazemos atribuições diretas de estados mentais a entidades sociais: “O Google *deseja* Feliz Natal com Doodle comemorativo”; “A NASA *pretende* enviar o primeiro ser humano a Marte na década de 2030”; “O partido *teme* o crescimento da oposição após denúncia de caso de corrupção”.

Apesar do uso cotidiano de sentenças de atribuição com entidades sociais figurando como seus sujeitos, quando paramos para refletir um pouco sobre esse tipo de sentença, nossa intuição é de que algo não está certo. Não parece ser o caso que instituições, organizações, grupos de amigos e famílias possuem, literalmente, uma mente capaz de sustentar estados mentais. O objetivo desse texto é revisar essa posição intuitiva. Nas páginas seguintes vou apresentar teorias contemporâneas sobre a natureza da mente que viabilizam a tese da mente de grupos seguindo algumas das propostas mais proeminentes de defesa dessa tese que encontramos na literatura.

1. Por que ser realista sobre mente de grupos?

Antes de nos debruçarmos sobre as questões mais técnicas, sobre metafísica da mente, creio que seja importante motivar a discussão sobre a tese da mente de grupos. Particularmente, tomo como relevante para a defesa de uma posição realista sobre mente de grupos o ponto geral apresentado na introdução acerca do uso de sentenças de atribuição de estados mentais ou sentenças de atribuição de outras propriedades que pressupõem estados mentais para grupos em virtude da importância que esse tipo de sentença de atribuição tem para as teorias filosóficas. Mais diretamente, realismo sobre mentes de grupos é uma condição necessária para se fazer Filosofia Social significativa.

Grande parte do empreendimento filosófico (em especial nas áreas da Epistemologia e da Ética) está centrado na análise de sentenças como¹:

1 Onde "S" é uma variável para sujeitos ou agentes; "p" é uma variável para proposições; "Φ" e "Ψ" são variáveis para ações. Ex: "João sabe que o Brasil está no Hemisfério Sul"; "Maria lançou a bola intencionalmente"; "Ricardo lançou a bola com a intenção de fazer a cesta"; etc.

“S sabe que p”;

“S está justificado a crer que p”;

“S Φ intencionalmente”;

“S Ψ com a intenção de Φ ”.

“S deve Φ ”;

“S Φ virtuosamente.”;

“S é culpado por uma ação Φ apenas se S for culpável pela ignorância de que Φ é moralmente errado.”

Um subconjunto particular de sentenças de atribuição que interessa à Filosofia diz respeito à atribuição de atitudes proposicionais (crença, desejo, intenção, dúvida, medo, confiança, surpresa). É óbvio que esse subconjunto de atribuições pressupõem uma mente, pois atitudes proposicionais são um tipo de propriedade mental.² Entretanto, é importante notar que a investigação de atitudes proposicionais não é isolada, essa investigação é relevante para a explicação, predição e racionalização de nós mesmos e de outros (Lindeman, 2024). Desta forma, outro subconjunto de atribuições que se mostram de interesse para a Filosofia (ação, moralidade, responsabilidade, justificação, conhecimento, racionalidade) pressupõe atitudes proposicionais e, conseqüentemente, pressupõe mente.

² Atitudes proposicionais são os exemplos paradigmáticos de atitudes intencionais. *Intencionalidade* é um termo técnico em Filosofia da Mente que busca expressar a capacidade que estados mentais têm de representarem objetos, propriedades ou estados de coisas (em inglês se caracteriza essa capacidade como “aboutness”, aproximadamente “ser sobre algo” [Jacob, 2023]). A posição padrão em Filosofia da Mente é que a maioria de nossos estados mentais têm essa capacidade: crenças e desejos são sobre; imaginações são sobre alguma coisa, podendo ela não existir, ser uma entidade ficcional.

Atitudes proposicionais formam um conjunto de propriedades mentais que possuem proposições como conteúdo. Para que a mente se relacione com algo e, portanto, tenha determinado conteúdo, é preciso que ela disponha da capacidade intencional. Assim fica bastante claro como atitudes proposicionais e intencionalidade estão relacionadas.

A tradição filosófica é marcada por um viés antropocêntrico. Os exemplos paradigmáticos de entidades que satisfazem o lugar de sujeito (S) das sentenças de atribuição que interessam à Filosofia são seres humanos, indivíduos da nossa espécie. Podemos derivar uma tese bastante restritiva a partir desse viés antropocêntrico:

*Individualismo de Sujeitos*³:

Apenas indivíduos da espécie humana podem ser sujeitos legítimos de sentenças de atribuição de estados mentais.

Um certo antropocentrismo enquanto ferramenta metodológica para a investigação me parece bastante razoável. Abordar questões sobre mente, consciência e estados mentais a partir da introspecção e, portanto, da experiência individual que cada um de nós (seres humanos) têm, me parece ser o primeiro passo inevitável para a investigação de entidades mentais. Apesar dessa concessão, uma tese como o *Individualismo de Sujeitos* não parece autorizada. O privilégio de acesso que temos às nossas próprias experiências conscientes não implica na negação da existência de outras mentes. Para começar, o *Individualismo de Sujeitos* já não se acomoda nos limites epistêmicos da introspecção. Ele concede legitimidade à atribuição de estados mentais a indivíduos cujas mentes não temos acesso privilegiado por introspecção. Eu não sei como é ser um astronauta, como é ser uma mãe ou como é ser meu colega de Departamento. A minha experiência individual não captura evidências diretas sobre aspectos mentais dos seres humanos que estão à minha volta e menos ainda de seres humanos com os quais não tenho qualquer contato.

³ Tradicionalmente, a literatura se refere a Quinton (1975) como um defensor dessa tese em discussões sobre o contexto social. Veja Ruivo (2017) para uma discussão crítica sobre essa interpretação.

Cichoski e Ruivo (2017) apresentam a tese de modo negativo: "Grupos não são sujeitos legítimos de sentença de atribuição de estados cognitivos." Os autores identificam em Tollefsen (2002) a discussão de uma tese semelhante, mas mais localizada, focada exclusivamente em propriedades epistêmicas.

A partir dessa constatação, podemos pensar também que, mesmo que não consigamos responder adequadamente à questão de como é ser um morcego, uma vaca ou um peixe (i.e. mesmo que não sejamos capazes de caracterizar as experiências fenomenológicas da consciência de uma outra espécie animal [Nagel, 2005]), o *Individualismo de Sujeitos* não deve ser uma razão para descartarmos que indivíduos dessas outras espécies possam figurar como sujeitos legítimos de atribuições de estados mentais. A minha incapacidade de caracterizar a sua experiência fenomenológica ao ler este texto não deveria invalidar a tese do *Individualismo de Sujeitos* e minha consequente legítima atribuição de crença (ou experiência visual) a você. Se este é o caso, isso mostra que a tese do *Individualismo de Sujeitos* é inadequadamente enviesada: liberal demais no que diz respeito a evidências de que seres humanos sejam sujeitos legítimos de atribuição; enquanto é restritiva demais para a extensão dessa legitimidade a outros tipos de entidades.

Agora, como definir o limite do conjunto dos sujeitos legítimos para atribuição de estados mentais? Quais são os critérios que uma entidade precisa satisfazer para integrar esse conjunto? Em grande medida, o presente texto se destina a apresentar algumas teorias sobre a mente que apresentam critérios para a resposta dessas questões que são compatíveis com a tese de que grupos tenham mentes. As teorias que tratarei a seguir são: funcionalismo (seção 3), mente estendida (seção 4), cognição (socialmente) distribuída (seção 5).⁴

⁴Originalmente, o texto também apresentava uma proposta pampsiquista baseada em Roelofs (2019) e Fonseca (2022, 2023). Infelizmente a complexidade da teoria e a limitação da extensão do texto impediram que essa proposta, com características bastante peculiares, bastante distintas das demais, pudesse ser discutida a contento neste momento. No entanto, ficam as indicações de obras sobre o pampsiquismo que abordam de que forma essa teoria sobre a natureza da mente também pode ser compatível com a tese de que grupos tenham mentes.

Entretanto, antes de avançarmos para a apresentação dessas teorias e como elas podem ser empregadas na defesa da tese da mente de grupos, mais precisa ser dito sobre a necessidade de negar o *Individualismo de Sujeitos* e, particularmente, sobre a concessão de que o sujeito (S) de sentenças de atribuição como aquelas apresentadas anteriormente possam ser grupos. Gostaria de explorar mais detalhadamente a razão apresentada anteriormente: que não é possível desenvolver uma Filosofia Social significativa sem esses recursos. A ideia aqui é que a “coletivização” de teorias filosóficas (epistêmicas, éticas, estéticas, etc.) depende da adoção de uma tese realista sobre mente de grupos. De modo muito direto, se grupos não podem ser sujeitos de sentenças como “S está justificado a crer que p”, “S deve Φ ”, etc., não é possível que esse tipo de entidade seja contemplado em teorias epistêmicas, éticas, estéticas, etc., padrão.

Talvez isso não seja um problema. Talvez grupos não sejam, realmente, sujeitos legítimos das atribuições necessárias para se encaixar nas teorias filosóficas padrão. Duas opções alternativas à necessidade do realismo sobre mente de grupos para a Filosofia Social são possíveis: (i) a Filosofia Social deve ser construída sem adesão às teorias filosóficas padrão, isto é, propriedades éticas, epistêmicas, estéticas, etc. de grupos são *sui generis*, não são do mesmo tipo que propriedades individuais; ou (ii) a Filosofia Social é um empreendimento reducionista, isto é, devemos buscar uma explicação detalhada da inter-relação entre indivíduos que pertençam a grupos para identificarmos que as propriedades relevantes (mentais, morais, epistêmicas) são sempre atribuíveis aos indivíduos, assim, a Filosofia Social não é uma investigação sobre grupos, mas sobre indivíduos e suas relações.

Particularmente, não creio que esses sejam os caminhos mais adequados para o desenvolvimento da Filosofia Social. Em primeiro lugar, se a motivação para essas linhas alternativas de investigação é a

contraintuitividade da tese da mente de grupos, o objetivo deste texto é mostrar que essa tese não é tão inadmissível, estando de acordo com teorias da Filosofia da Mente atuais. Em segundo lugar, parece muito mais custoso (i) partir do zero, desenvolver novas teorias sem considerar as bases das teorias filosóficas padrão ou (ii) ignorar atribuições do nível do grupo em favor de uma investigação pormenorizada de relações individuais.

Justamente adotando uma posição contrária à estratégia reducionista, podemos encontrar um argumento a favor da tese da mente de grupos a partir da prática científica no campo das Ciências Humanas e Sociais. Áreas como Psicologia Social, Economia, Administração, Direito, Sociologia, dentre outras, fazem uso de sentenças de atribuição de estados mentais e ações a grupos rotineiramente. Uma postura naturalista orientada pelo compromisso ontológico pode sustentar um argumento a favor da tese da mente de grupos da seguinte forma (adaptado de List, 2016, p. 4):

- (1) Nossas melhores teorias científico-sociais de um certo fenômeno social - p.e. nossas melhores teorias sobre o comportamento das firmas no mercado - atribuem agência baseada em crenças e desejos para os coletivos envolvidos, geralmente ao representá-los como agentes no sentido das Teorias da Decisão e dos Jogos.
- (2) Nós devemos, ao menos falivelmente, adotar os compromissos ontológicos de nossas melhores teorias científicas em qualquer domínio.
- (C) Nós devemos, ao menos falivelmente, adotar o compromisso ontológico das nossas melhores teorias científico-sociais com relação a coletivos que possuem agência baseada em crenças e desejos.

Percebemos assim que a motivação geral para a defesa da tese da mente de grupos está ancorada no compromisso ontológico. As teorias científicas das Ciências Humanas e Sociais e as teorias filosóficas da Filosofia Social fazem uso de atribuição de estados mentais a grupos.⁵ Como o argumento acima indica, nossa posição inicial (falível, derrotável) deveria ser pela interpretação literal dessas sentenças de atribuição. Uma posição inicial baseada em uma intuição antirrealista parece mal motivada.

2. A intuição antirrealista sobre mente de grupos

Talvez eu tenha me apressado. Talvez a posição antirrealista seja motivada. Nesta seção apresentarei, brevemente, pressupostos metafísicos sobre a natureza da mente que seriam contrários à tese da mente de grupos. Com certeza as nossas intuições não dão suporte à tese de que grupos tenham ou possam ter mentes. Concedo que a tese de que exista, literalmente, uma mente de grupos seja contraintuitiva. Há uma intuição muito forte de que nossos familiares, amigos, colegas e demais seres humanos possuem estados mentais; uma intuição favorável à tese de que animais não humanos (pelo menos aqueles com um sistema nervoso semelhante ao dos seres humanos) também possuem estados mentais. Entretanto, há uma intuição forte contra a tese de que sistemas de latinhas de cerveja arrançadas com fios e grupos organizados de pessoas possuem estados mentais (Schwitzgebel, 2015). Vejamos alguns pressupostos metafísicos sobre a mente que podem

⁵ Margaret Gilbert (1989; 2002) parte de atribuições cotidianas de estados mentais a grupos para motivar sua posição realista. Concordo com Gilbert que esse é um ótimo primeiro passo evidencial e de motivação para a investigação do tema. Interpreto que essas atribuições possuem uma função metodológica no trabalho da autora. Entretanto, essa função argumentativa que atribuições cotidianas poderiam ter me parece um pouco mais delicada, pois um argumento via compromisso ontológico da linguagem natural é controverso (Bricker, 2014).

servir de base para essas intuições sobre quais entidades possuem estados mentais e quais não possuem estados mentais.

À primeira vista, o fisicalismo, a tese de que toda a realidade é física⁶, poderia sustentar uma objeção direta à tese da mente de grupo. Um fisicalismo mais radical poderia ser uma ameaça a qualquer teoria sobre mentes: se mentes (individuais ou coletivas) são um tipo de entidade não física, então elas não existem (Ramsey, 2019). Uma tese fisicalista um pouco mais fraca poderia defender que entidades mentais existem, mas estão relacionadas com entidades físicas. Diferentes propostas de dependência ontológica poderiam ser apresentadas para a explicação dessa relação. Uma dependência mais rígida poderia ser sustentada pela tese da identidade entre certos tipos de entidades mentais e físicas (por exemplo, a dor é o estímulo das fibras nervosas C); teses de dependência menos reducionistas poderiam ser defendidas por meio de relações de superveniência ou realização (propriedades mentais são supervenientes a propriedades físicas).

Essas versões mais fracas de fisicalismo parecem embasar as pesquisas das ciências cognitivas. Essas teses de dependência de entidades mentais sobre entidades físicas motiva a investigação de correlações entre (relatos dos sujeitos experimentais sobre seus) estados mentais e (o rastreamento de) atividades em partes do cérebro ou do sistema nervoso. E aqui chegamos a um pressuposto metafísico sobre a natureza da mente, bastante aceito hoje em dia, que invalidaria qualquer tese sobre mente de grupos: a mente depende do cérebro. Podemos estruturar essa posição em um argumento da seguinte forma:

⁶ Ou depende de entidades físicas. Para uma discussão mais precisa das variantes do fisicalismo e sua relação com aspectos metafísicos da Filosofia da Mente, veja Gouvea (2020).

Argumento FISCALISTA contra mente de grupos:

- (1) A mente depende do cérebro.
- (2) Grupos não têm cérebros.
- (C) Grupos não têm mentes.

Além do argumento fiscalista, outro pressuposto metafísico sobre a natureza da mente que invalidaria qualquer tese sobre mente de grupos diz respeito à fenomenologia. Mais especificamente, a intuição de que experiências fenomênicas são um traço fundamental e, portanto, necessário de uma mente. “Aspecto qualitativo”, “perspectiva de primeira pessoa”, “conteúdo fenomênico” ou o “como é ser (um morcego)” são expressões que tentam elucidar essa característica peculiar de experimentar a posse ou ocorrência de um estado mental. Esse é um aspecto tão relevante para a nossa compreensão da mente que, intuitivamente, parece se apresentar como uma condição necessária para sua existência. De modo bastante similar à abordagem fiscalista, agora teríamos a afirmação de que a mente depende de consciência e isso, aparentemente, seria um obstáculo para qualquer tese realista sobre mente de grupos. Podemos estruturar essa posição em um argumento da seguinte forma:

Argumento da CONSCIÊNCIA contra mente de grupos:⁷

- (1) A mente depende da consciência.
- (2) Grupos não têm consciência.
- (C) Grupos não têm mentes.

⁷ Szanto (2014, p. 100, tradução nossa) apresenta um argumento muito similar:

- (1) Ter uma mente é ter estados mentais (auto) conscientes com conteúdo fenomênico.
- (2) Grupo não têm estados mentais (auto) conscientes com conteúdo fenomênico.
- (3) Portanto, grupos não têm mentes.

O objetivo desse texto não será de desarmar esses argumentos. Minha abordagem sobre respostas a esses desafios será mais indireta. Nas próximas seções apresentarei propostas teóricas sobre a natureza da mente que tornam plausível a tese da mente de grupos. Ao invés de empregar uma abordagem negativa, buscando a refutação das intuições apresentadas nessa seção, vou apresentar teorias que parecem ser compatíveis com a existência de mentes de grupo.

Antes de avançar, no entanto, enfatizo que dentre as premissas dos argumentos apresentados nessa seção certamente “Grupos não têm cérebros” é inquestionável. As demais (“A mente depende do cérebro”; “A mente depende da consciência”; “Grupos não têm consciência”) não são consensuais na literatura. Ainda que “Grupos não têm consciência” tenha um histórico de endosso (Szanto, 2014; Tollefsen, 2015; List, 2016), existem propostas que buscam questioná-la (Schwitzgebel, 2015; Kramer, 2021). Além disso, se “Grupos não têm cérebros” é uma tese consensual, todas as posições que serão apresentadas a seguir deverão negar a tese de que “A mente depende do cérebro” para viabilizar a tese da mente de grupos.

3. Funcionalismo

Estratégias funcionalistas foram largamente empregadas para tratar de questões relativas à responsabilidade coletiva desde a década de 1960 (May e Hoffman, 1992; Smiley, 2022). Particularmente, a proposta de Peter French (1979; 1984; 1995), bastante influente nessa temática específica, se concentra na tese de que corporações e outras organizações formais possuem estruturas de tomada de decisão que as capacitam a serem tomadas como pessoas morais, isto é, como entidades capazes de agir intencionalmente e, portanto, serem

responsabilizadas moralmente. Essa tese foi significativamente aperfeiçoada por List e Pettit (2011).

Antes de especificar os detalhes dessas propostas que tratam diretamente da questão da agência no contexto social, será necessário esclarecer um pouco mais a tese funcionalista sobre a mente adotada por suas versões mais sofisticadas. A tese funcionalista sobre a mente visa renegociar uma de nossas intuições metafísicas sobre a mente. Como o nome da abordagem já explícita, a tese funcionalista defende que mentes são entidades funcionais. Isso quer dizer que a natureza da mente (o que caracteriza o que é ser uma mente, a propriedade fundamental de algo que é uma mente) é relativa ao desempenho de uma função e não relativa a sua constituição material (Polger, 2024).

A apresentação de outras entidades funcionais pode elucidar esse pressuposto metafísico. Em geral, artefatos são entidades funcionais. Cadeiras e ratoeiras são exemplos desse tipo de artefato. Um objeto só é uma cadeira se satisfaz o critério funcional de poder sustentar o corpo de um ser humano mediano sentado. Um objeto só é uma ratoeira se satisfaz o critério funcional de capturar ratos. O ponto importante desses exemplos é notar que cadeiras e ratoeiras podem diferir no que diz respeito a sua constituição material, mas são cadeiras e ratoeiras por serem objetos que satisfazem os critérios funcionais apresentados. Cadeiras podem ser compostas de madeira, plástico, pedra, concreto, etc. Ratoeiras podem ser físicas ou químicas, podem ser gaiolas que aprisionam ratos, compostos químicos (adesivos) que prendem ratos, mecanismo com guilhotina; até mesmo mecanismos com ganchos e armas de fogo já foram propostos para capturar e eliminar esses animais. Esse é o ponto relevante que ilustra a “renegociação metafísica” proposta pela tese funcionalista da mente. Se a mente é uma entidade funcional, assim como cadeiras e ratoeiras, não é em virtude

de sua composição material (orgânica, neuronal, silício, éter, eletricidade, magnetismo), mas sim em virtude das funções que desempenha que ela é aquilo que é.

Mas quais funções seriam características da mente? As abordagens coletivas citadas anteriormente vão propor uma espécie de behaviorismo funcionalista. As características funcionais relevantes para a identificação de algo operando em virtude de possuir uma mente seriam evidenciadas por determinado tipo de comportamento.⁸ Desta forma, a função característica de algo que possui uma mente terá entradas (*inputs*) e saídas (*outputs*) que descrevem o comportamento próprio de um agente. Cabe enfatizar que não é qualquer tipo de comportamento ou qualquer tipo de perfil funcional (conjunto de entradas e saídas típicas) que caracterizaria um agente. O termo “agente” aqui é entendido em um sentido um pouco mais estrito (mas é a concepção tradicional em Filosofia da Ação [veja Schlosser, 2019]), pois não é qualquer entidade que participa de uma relação causal, mas sim

⁸ Strohmaier (2020) argumenta que List & Pettit (2011) seriam mais apropriadamente considerados interpretativistas do que funcionalistas. Como o autor destaca, a diferença substancial entre essas duas perspectivas se dá na especificação da função (conjunto de entradas e saídas) estabelecida pela teoria. Interpretativistas são mais behavioristas, destacam comportamentos externos (pegar um guarda-chuva, condenar o réu, organizar os livros da biblioteca) como as saídas típicas da função característica de um sistema ou criatura que possui uma mente. Por outro lado, funcionalistas destacam papéis funcionais de mecanismos internos (os estados mentais específicos: crenças, desejos, intenções) do sistema ou criatura (a intenção de pegar o guarda-chuva produz determinados movimentos corpóreos do agente, a decisão de que o réu é culpado resulta na condenação, o plano para organização da biblioteca orienta essa atividade).

Uma observação similar pode ser encontrada na crítica de Dahan (2017) a Schwitzgebel (2015). A distinção é a mesma. Dahan afirma que Schwitzgebel não apresenta propriamente um funcionalismo em sua argumentação, pois as funções ou papéis causais identificados por ele são propriedades do sistema ou criatura (ser orientado por objetivos, se autoprotger e autopreservar de modo reflexivo, responder inteligentemente a oportunidades e ameaças, ser capaz de se comunicar com outras entidades semelhantes). A rigor, para um funcionalista a característica que torna algo um estado mental deveria dizer respeito ao papel causal que esse algo (uma parte do sistema) tem no sistema. Assim, uma análise funcionalista deveria destacar os itens funcionais das subunidades de um sistema (estados mentais específicos como crença, desejo, dor, etc.) ao invés do comportamento externo desse sistema (a execução de determinada ação).

aquela entidade que é capaz de produzir eventos em virtude de uma estrutura funcional específica, caracterizada pela presença de estados mentais que tornam essas ações em ações intencionais.⁹

French (1979; 1995) busca argumentar que organizações estruturadas (paradigmaticamente empresas e corporações) possuem três características que evidenciam sua capacidade de executar ações intencionais: (1) elas agem intencionalmente, no sentido de que suas ações são guiadas por um propósito; (2) elas são racionais, pois estabelecem seus propósitos e parecem ser racionalmente responsáveis a circunstâncias relacionadas a esses propósitos; (3) elas são capazes de responder a críticas, elas ajustam seu comportamento ou oferecem razões em resposta a argumentos. Esse conjunto de comportamentos é típico de um agente intencional (pessoa moral), cumprindo assim a satisfação de critérios funcionais para a atribuição de estados mentais a esse tipo de entidade.¹⁰

⁹ É essa característica funcional envolvida nas ações intencionais, a capacidade de agência, que é tomada como sendo relevante para a atribuição de responsabilidade moral (atribuições de culpa ou elogio): “O julgamento de que uma pessoa é moralmente responsável por seus comportamentos envolve - ao menos em uma primeira aproximação - a atribuição de certos poderes e capacidades para aquela pessoa e a constatação de que seus comportamentos são causados (da maneira correta) a partir do fato de que a pessoa possui, e exerceu, esses poderes e capacidades. Seja qual for a teoria correta sobre os poderes e capacidades em questão, sua posse qualifica um agente como sendo responsável moralmente em um sentido geral: isto é, como alguém que pode ser responsável moralmente por determinados exercícios de agência. Seres humanos adultos normais podem possuir os poderes e capacidades em questão e animais não humanos, crianças muito pequenas e aqueles que sofrem de deficiências graves de desenvolvimento ou demência (para dar alguns exemplos) são geralmente considerados como não os possuindo” (Talbert, 2024, tradução nossa).

¹⁰ É importante salientar que French (1995) faz ressalvas sobre a atribuição de estados mentais a organizações estruturadas; ele não é tão realista quanto nossa descrição faz parecer. Sua proposta busca defender a possibilidade de que um tipo de grupo (organizações estruturadas) pode ser um agente intencional (em virtude de satisfazer os critérios funcionais discutidos acima) e, portanto, pode ser uma pessoa moral, ou seja, está sujeito a atribuição de responsabilidade. No entanto, não parece muito coerente manter uma restrição relativa a estados mentais quando se concede que uma determinada entidade é um agente intencional (Tollefsen, 2015).

Essa seria uma posição bastante heterodoxa em Filosofia da Ação visto que estados mentais são tomados como necessários para a execução de uma ação intencional: “[A teoria padrão] diz, muito grosseiramente, que algo é uma ação intencional e foi realizado por razões, apenas caso isso seja causado pelos estados e eventos mentais corretos e da maneira correta.” (Schlosser, 2019, tradução

Também explorando critérios para a identificação da capacidade de agência, List e Pettit (2011) destacam características básicas que uma entidade deve instanciar para ser considerada um agente: (1) ela deve ter estados representacionais que retratam como as coisas são no ambiente; (2) ela deve ter estados motivacionais que especificam como o agente gostaria que as coisas do ambiente fossem; (3) ela deve ter a capacidade de processar estados representacionais e motivacionais, conduzindo intervenções no ambiente quando esse ambiente não estiver de acordo com as especificações motivacionais. Os autores combinam esses critérios em uma definição bastante geral de agente: um sistema que apresenta estados representacionais, estados motivacionais, a capacidade de processar esses estados, bem como de agir com base neles. Note que essa abordagem é explicitamente mentalista-cognitivista. As características centrais de um agente dizem respeito a estados (mentais) representacionais e motivacionais (crenças e desejos, respectivamente), seu processamento (raciocínio) e a execução de ações condizentes com eles (satisfação de condições funcionalistas).

Tanto French quanto List & Pettit argumentam que (alguns) grupos podem satisfazer os requisitos de agência por eles discutidos em virtude do modo como estão organizados, isto é, em virtude de uma estrutura organizacional específica que possibilita a ocorrência de processos de tomadas de decisão racionais, processos que exibem o nível de racionalidade característico de um agente. A argumentação de French (1979) é baseada na caracterização de uma estrutura interna de decisão.

nossa). Esse ponto pode, inclusive, suprir a lacuna entre perfis funcionais de agentes ou sistemas e perfis funcionais de estados mentais ou subsistemas que motiva a crítica da nota 6. Ainda que a descrição funcional esteja no nível do sistema, é em virtude de pressupostos amplamente aceitos pela Filosofia da Ação que se deveria supor o exercício de papéis funcionais de estados mentais sendo instanciados em circunstâncias nas quais uma ação do sistema esteja sendo executada. List & Pettit (2011) e Tollefsen (2015) desenvolvem esse argumento abraçando o interpretativismo baseado na teoria da postura intencional de Daniel Dennett (1987).

Essa estrutura é organizada por meio de (i) uma hierarquia que ordena os setores e níveis da organização determinando as relações de comando e poder e (ii) uma diretriz corporativa, um conjunto de regras para a tomada de decisão. A estrutura é ativada por indivíduos (seres humanos particulares) que ocupam as posições da hierarquia.

List & Pettit (2011) desenvolvem uma abordagem bastante técnica do processo de tomada de decisão em grupos estruturados por meio de ferramentas teóricas da teoria da escolha social.¹¹ Esse aparato permite aos autores apresentar casos onde a posição do grupo pode diferir da posição individual de todos os seus membros. Esse tipo de resultado é relevante para uma posição realista sobre mentes de grupos, pois exemplifica propriedades mentais coletivas que são irredutíveis às propriedades mentais de seus membros individuais.

A proposta dos autores é desenvolvida a partir de casos de Dilema Discursivo. Esses são casos onde “votações majoritárias em proposições interconectadas podem levar a julgamentos de grupo inconsistentes, mesmo quando os julgamentos individuais são completamente consistentes” (List; Pettit, 2011, p. 46, tradução nossa).¹² Essa caracterização geral do Dilema Discursivo não é muito clara, mas um exemplo pode nos auxiliar a identificar uma posição inconsistente por parte do grupo, mesmo que todos os seus membros tenham posições consistentes. Vejamos um desses exemplos a seguir:

¹¹ List (2022, tradução nossa) define esse campo de investigação da seguinte forma: “A teoria da escolha social é o estudo dos procedimentos e mecanismos de decisão coletiva. Ela não é uma única teoria, mas um conjunto de modelos e produtos relativos à agregação de entradas (*inputs*) individuais (e.g. votos, preferências, julgamentos, bem-estar) em saídas (*outputs*) coletivas (e.g. votos, preferências, julgamentos, bem-estar).”

¹² Essa definição de Dilema Discursivo é uma generalização de um problema particular da literatura jurídica sobre decisões colegiadas que ficou conhecida como Paradoxo Doutrinal (Kornhauser; Sager, 1986). A generalização está baseada na identificação de que o problema surge por conta de haver um conjunto de proposições logicamente interconectadas sendo considerado pelo processo de agregação. No Paradoxo Doutrinal essa conexão era estabelecida pela Doutrina Jurídica.

Índex da Democracia (adaptado de Cichoski; Ruivo, 2015)

Analistas estão interessados em avaliar o nível de democracia dos países ao redor do mundo. Para realizar essa avaliação, eles estabelecem que para ser uma democracia plena o país deveria satisfazer três condições: (i) ter processo eleitoral democrático, (ii) ter funcionamento do governo condizente com a democracia e (iii) ter alto nível de participação política.

Podemos identificar um argumento envolvido nessa forma de estruturar a análise do nível de democracia de países. Isso revela que a análise oferece um conjunto interconectado de proposições da seguinte forma:

P - O País X tem um processo eleitoral democrático;

Q - O governo do País X tem funcionamento condizente com a democracia;

R - O País X tem alto nível de participação política;

C - O País X é uma democracia plena.

A interconexão lógica nesse argumento pode ser expressa da seguinte forma: $P \wedge Q \wedge R \rightarrow C$.

Agora, suponha que há um grupo de analistas averiguando o caso específico do Brasil e eles chegam a essas posições individuais:

Analista 1 - crê que o Brasil não é uma democracia plena, pois considera que não há um processo eleitoral democrático no Brasil, ainda que considere o funcionamento do governo e a participação política condizentes com as de uma democracia plena;

Analista 2 - crê que o Brasil não é uma democracia plena, pois considera que o funcionamento do governo não é adequado para uma democracia plena, ainda que considere o processo eleitoral democrático e que há alto nível de participação política;

Analista 3 - crê que o Brasil não é uma democracia plena, pois não considera a participação política suficientemente alta, ainda que considere o processo eleitoral e o funcionamento do governo adequados para uma democracia plena.

A partir desse posicionamento individual de cada analista, vejamos o que aconteceria se a decisão final do grupo a respeito de cada um dessas proposições fosse resultado de uma agregação por maioria, i.e. a posição do grupo é aquela que for a maioria das posições individuais. A tabela a seguir ilustra a distribuição de posições individuais e o resultado da agregação:

	P	Q	R	C
Analista 1	Falsa	Verdadeira	Verdadeira	Falsa
Analista 2	Verdadeira	Falsa	Verdadeira	Falsa
Analista 3	Verdadeira	Verdadeira	Falsa	Falsa
Grupo (Maioria)	Verdadeira	Verdadeira	Verdadeira	Falsa

Tabela 1. Agregação por maioria no caso *Índex da Democracia*.

O caso *Índex da Democracia* é uma instância do Dilema Discursivo. A posição do grupo é inconsistente, enquanto a posição de cada indivíduo é consistente. Todos os analistas estão, individualmente, seguindo a restrição imposta pela interconexão lógica estabelecida pelas proposições sendo consideradas. Como vimos antes, a relação $P \wedge$

$Q \wedge R \rightarrow C$ identifica um conjunto de condições conjuntamente necessárias para que a proposição C seja verdadeira. Na perspectiva individual, cada um dos analistas julgou que uma dessas condições necessárias não foi satisfeita e, portanto, sustentam que a proposição C seja falsa. Desta forma, cada um deles sustenta um conjunto consistente de posições sobre essas proposições.

O problema está na última linha da tabela, no resultado da agregação por maioria, ou seja, na posição final que o grupo vai assumir. Essa linha apresenta um conjunto inconsistente de posições. Foi estabelecido pelo caso que haveria uma restrição na interconexão lógica expressa por $P \wedge Q \wedge R \rightarrow C$. Essa restrição não está sendo respeitada pela posição final do grupo. Como pode ser observado na última linha da tabela, o grupo considera que foram satisfeitas as condições necessárias, mas nega a conclusão que deveria se seguir disso. A posição do grupo seria algo como: “O Brasil atende a todas as condições necessárias para ser uma democracia plena, portanto, não consideramos que o Brasil seja uma democracia plena”.

List e Pettit não acreditam que grupos estejam organizados com esse tipo de fragilidade em sua racionalidade. Eles discutem que o tratamento de conjuntos de proposições interconectadas demandam funções de agregação especiais. A agregação por simples maioria gera Dilemas Discursivos, então não é esse o meio pelo qual decisões coletivas são tomadas quando se trata de conjuntos de proposições interconectadas. A interpretação dessa interconexão lógica entre proposições enquanto um argumento ilustra as opções disponíveis para funções de agregação especiais: podemos agregar pelas premissas ou pela conclusão. Voltando ao caso do *Índex da Democracia*, caso a agregação se dê pela conclusão, a maioria dos analistas julgou que o Brasil não era uma democracia plena (C é falsa). Portanto, essa seria a

posição final do grupo, mas só essa proposição será considerada pelo grupo, ele não teria posição sobre as demais proposições. Entretanto, caso se adote a agregação pelas premissas, o grupo aceita que o Brasil satisfaz as condições necessárias para ser uma democracia plena, visto que a maioria dos analistas julga que P, Q e R são verdadeiras, e infere, em virtude da interconexão lógica $P \wedge Q \wedge R \rightarrow C$, que o Brasil é uma democracia plena.

Duas consequências importantes podem ser observadas pela adoção da agregação pelas premissas. (1) Essa modalidade de agregação, de modo geral, exemplificaria um processo literal de raciocínio de grupo. É o grupo que deriva a conclusão a partir da agregação das premissas e da consideração da restrição expressa pela interconexão lógica das proposições envolvidas. Esse resultado permite a defesa da plausibilidade de que grupos possuem uma mente autônoma e possam ser consideradas pessoas institucionais (Pettit, 2003). (2) Mais particular ao tipo de caso do nosso exemplo, essa é uma instância onde o grupo tem uma posição que não é sustentada por nenhum de seus membros individuais. Ao adotar a agregação pelas premissas no caso *Índex da Democracia*, deveríamos modificar a atribuição final do grupo com relação à proposição C. Na tabela você poderá conferir que nenhum dos membros individuais considera a proposição C verdadeira, enquanto o grupo deveria adotar essa posição se adotasse essa modalidade de agregação.¹³ (1) e (2) configuram evidências em favor da tese da mente

¹³Tecnicamente, os autores destacam que há uma superveniência holística quando se adota uma função de agregação pelas premissas em conjuntos de proposições logicamente interconectadas. A superveniência holística afirma que "o conjunto de atitudes de grupo sobre [o conjunto interconectado de] proposições é determinado pelos conjuntos de atitudes individuais sobre essas proposições" (List; Pettit, 2011, p. 69, tradução nossa). Esse resultado confere autonomia ao nível coletivo, onde a posição do grupo pode diferir das posições individuais de seus membros, mas explica como o nível individual é a base de superveniência desse resultado. É a posição dos membros individuais (sobre as premissas) que determina a posição do grupo (tanto sobre as premissas, quanto sobre a conclusão). Assim, há uma

de grupos. (1) identifica processos cognitivos que são próprios do grupo e (2) identifica a possibilidade de que o grupo tenha uma posição que não seja adotada por nenhum de seus membros individuais.

4. Mente estendida

Um segundo caminho que a Filosofia da Mente contemporânea oferece para a plausibilidade da tese da mente de grupos é pavimentado pela tese da mente estendida. A tese da mente estendida defende que a mente não está confinada dentro de nossos cérebros. Mais audaciosamente, é a ideia de que a mente se estende para além de nossos corpos biológicos acomodando outras entidades do ambiente no qual estamos inseridos. As entidades ambientais mais representativas do processo de extensão da mente são artefatos como calculadoras, computadores e smartphones. Mas simples anotações em cadernos ou notas autoadesivas também poderiam ser considerados mecanismos para a extensão da mente.

Pode até ser possível argumentar que a mente estendida não é um novo argumento, pois sua fundamentação está amparada em uma tese funcionalista: o Princípio da Paridade. O Princípio da Paridade costuma ser apresentado a partir da seguinte passagem do texto seminal da proposta:

Se, conforme nós confrontamos alguma tarefa, uma parte do mundo funciona como um processo, o qual, fosse ele executado na cabeça, nós não hesitaríamos em reconhecê-lo como parte do processo cognitivo, então essa parte do mundo é (assim alegamos) parte do processo cognitivo (Clark; Chalmers, 1998, p. 8¹⁴).

explicação metafísica muito parcimoniosa sobre como se dá o resultado de divergência entre níveis, a autonomia do nível coletivo e o surgimento de uma realidade social.

¹⁴ Estou utilizando a tradução apresentada por Eros Moreira Carvalho (2020). Veja o artigo do prof. Eros para críticas sobre a fundamentação da tese da mente estendida por meio do princípio da paridade.

Novamente temos os elementos centrais do funcionalismo: (i) a base material (orgânica, neuronal, silício, éter, eletricidade, magnetismo) de composição daquilo que faz parte do processo sendo analisado não é relevante e (ii) o aspecto relevante para identificar algo como sendo parte do processo cognitivo diz respeito à uma característica funcional, à sua participação na tarefa que está sendo realizada. O ponto adicional que a tese da mente estendida permite destacar é que não apenas a base material de composição da mente pode ser distinta, mas a localização desse componente material da cognição pode estar fora do crânio ou do corpo de uma criatura que possui uma mente. Uma parte do mundo, uma entidade do ambiente ou um artefato pode desempenhar algum papel funcional no processo cognitivo e, portanto, ser parte da mente de um sistema.

Clark e Chalmers (1998) apresentam alguns exemplos para elucidar a tese da mente estendida. Vou analisar um deles, o caso de Otto e seu caderno. Suponha que Otto possui algum tipo de condição de perda de memória (no caso original ele sofre de Alzheimer) e cria um mecanismo para atenuar essa situação: ele passa a anotar informações relevantes em um caderno. Com o tempo, o uso do caderno por Otto se torna tão natural que não seria um exagero dizer que Otto incorpora esse artefato¹⁵, em outras palavras: “Otto se torna fluente, habilidoso e confiável na

¹⁵ Merleau-Ponty (1999, p. 198-199) oferece uma ótima descrição de incorporação de instrumentos quando tenta esclarecer o que é um hábito: “A bengala do cego deixou de ser para ele um objeto, ela não mais é percebida por si mesma, sua extremidade transformou-se em zona sensível, ela aumenta a amplitude e o raio de ação do tocar, tornou-se o análogo de um olhar. Na exploração dos objetos, o comprimento da bengala não intervém expressamente e como meio-termo: o cego o conhece pela posição dos objetos, antes que a posição dos objetos por ele. A posição dos objetos está imediatamente dada pela amplitude do gesto que a alcança e no qual está compreendido, além da potência de extensão do braço, o raio de ação da bengala. (...) Habituar-se a um chapéu, a um automóvel ou a uma bengala é instalar-se neles ou, inversamente, fazê-los participar do caráter volumoso de nosso corpo próprio. O hábito exprime o poder que temos de dilatar nosso ser no mundo ou de mudar de existência anexando a nós novos instrumentos.”

recuperação de informações registradas no caderno” (Carvalho, 2020, p. 148). Essas características funcionais relacionadas ao resgate de informações do caderno por parte de Otto podem ser comparadas com as atividades cognitivas de um outro indivíduo, Inga, que utiliza uma estrutura exclusivamente biológica hidrocarbonada e neuronal (seu cérebro) para resgatar informações previamente registradas. Desta forma, pela satisfação do princípio da paridade, podemos concluir que o caderno de Otto faz parte do seu processo cognitivo de memória, assim como o cérebro de Inga desempenha a mesma função.

A aplicação desses processos cognitivos em uma instância específica de comportamento elucidada ainda mais a tese (adaptado de Clark e Chalmers, 1998). Considere agora que Inga ouve de sua colega de trabalho que está ocorrendo um evento especial para o dia dos namorados no Museu de Arte Moderna (MoMA). Ela forma o desejo de ir até o museu e resgata a crença de que ele fica na Rua 53. Seu desejo e sua crença são estados mentais que motivam e guiam sua ação de ir até o museu. Agora considere que Otto também deseje ir ao MoMA ver a coleção de “Construções Corporais” (*Body Constructs*). Seu problema de memória impede que ele consiga se lembrar do endereço do museu, mas ele vasculha seu caderno em busca dessa informação. Otto encontra o registro do endereço do MoMA em seu caderno e se dirige à Rua 53. Novamente, o funcionalismo por trás do princípio de paridade parece indicar que Inga e Otto realizaram os mesmos processos cognitivos e seu comportamento pode ser explicado pelos mesmos estados e processos mentais: (i) o desejo de ir ao MoMA; (ii) o processo de resgate da informação de onde o MoMA se localiza; (iii) a crença de que o MoMA fica na Rua 53.

Tollefsen (2006, p. 144, tradução nossa) sintetiza bem a ênfase que Clark e Chalmers conferem ao papel do funcionalismo na defesa da tese

da mente estendida, especificando que os autores apelam para uma interpretação de senso comum da dinâmica causal (perfil funcional) do comportamento dos sujeitos no caso Otto:

Da nossa perspectiva, enquanto intérpretes de comportamentos, Otto será guiado pela informação contida em seu caderno da mesma forma que ele seria guiado se essa informação estivesse em sua memória de curto prazo. Ele fará o mesmo tipo de coisas, dirá o mesmo tipo de coisas e formará os mesmos tipos de atitudes proposicionais com base na informação resgatada de seu caderno. Esse é um sentido frouxo de “dinâmica causal”, mas é, claramente, o sentido que incita a nossa prática cotidiana de prever e explicar o comportamento das pessoas. Eu posso prever com alguma precisão que, depois de olhar para seu caderno, Otto seguirá em direção ao museu.

Obviamente nem todas as entidades do ambiente em que uma criatura está inserida serão incorporadas em processos cognitivos, mesmo aquelas que tratam de registro, acesso e processamento de informações podem falhar em ser apropriadamente acopladas. Podemos usar a calculadora para fazer contas, mas provavelmente o fazemos de forma muito esporádica; uma agenda pode registrar todos nossos compromissos, mas ela pode ficar exclusivamente no escritório; você pode comprar um *smartwatch*, mas não saber utilizar suas funcionalidades, recorrendo ao aparelho apenas para ver as horas; usamos a Internet para acessar informações, mas não aceitamos (ao menos não deveríamos) tudo que ali encontramos acriticamente; também nos informamos pela imprensa (jornal, rádio, televisão), mas nosso contato com esses meios de comunicação é passivo (epistemicamente), tomamos testemunhos desses meios, não fazemos o trabalho de investigação necessário para justificar (em primeira mão) as informações ali registradas.

Clark e Chalmers (1998) destacam as seguintes características que esboçam o que estaria envolvido no devido acoplamento de uma criatura biológica com algum recurso externo: (1) constância - a criatura

recorrentemente utiliza o recurso para realizar o processo cognitivo; (2) fluidez - a criatura domina o uso do recurso; (3) endosso automático - as informações geradas pelo uso do recurso são tomadas acriticamente; (4) endosso prévio - no caso de registros de informações, como no caso Otto, as informações registradas no recurso externo foram previamente endossadas, sendo essa uma razão para terem sido registradas e automaticamente endossadas quando novamente acessadas.

Tollefsen (2006) utiliza esses quatro critérios para justificar o acoplamento de outros indivíduos. Seu argumento é construído a partir de uma variação do caso de Otto e seu caderno: Olaf e Olga. Considere que Olaf e Olga estejam casados há 30 anos. Olaf não sofre de Alzheimer, mas é bastante esquecido, tendo dificuldades em lembrar nomes de pessoas, compromissos, números de telefone, endereços etc. Olga, pelo contrário, tem uma memória de elefante e consegue se lembrar desses pequenos detalhes facilmente. Considere ainda que eles estão quase sempre juntos. Diante dessas características, parece que Olaf e Olga atendem aos critérios para acoplamento tanto quanto Otto e seu caderno: (1) constância - Olaf constantemente pede informações sobre compromissos, números de telefone e endereços a Olga; (2) fluidez - Olga é bastante responsiva às solicitações de Olaf, a comunicação entre eles é adequada e “enquanto uma parceira cognitiva afetuosa e comprometida, Olga está sempre disponível - na saúde, na doença e na perda de memória” (Tollefsen, 2006, p. 143, tradução nossa); (3) endosso automático - Olaf não questiona as informações transmitidas por Olga, inclusive a utiliza como fonte de verificação de coisas que ele lembra “Olga, é hoje o dia da consulta ao médico?”; (4) endosso prévio - Olga não é enxerida, ela respeita a privacidade de Olaf, e não inventa as informações que comunica a Olaf. Geralmente, Olaf comunica seus compromissos e solicita que Olga o ajude a lembrar.

O acoplamento de indivíduos parece autorizado pela tese da mente estendida¹⁶, mas ele gera alguns questionamentos. Uma questão central que surge da possibilidade de aplicação da tese da mente estendida para contextos coletivos é: será que Olga agora faz parte da mente de Olaf? Uma resposta para essa questão não é muito clara. Julgar os limites da mente estendida parece mais fácil nos exemplos usuais tratados pelos proponentes da tese, onde a extensão ocorre com o acoplamento de componentes materiais não biológicos do ambiente em que uma criatura biológica está inserida. Otto e seu caderno, você e seu celular, um motorista e seu GPS, etc. são exemplos que parecem acomodar essa noção de literal extensão da mente de um indivíduo. Talvez haja um viés antropocêntrico na prioridade ontológica da criatura biológica que compõe o sistema. A mente se estende, mas seu ponto de partida é a mente biológica. Harris (2020, p. 7, tradução nossa) apresenta essa interpretação da tese da mente estendida na seguinte proposta de definição:

Tese da mente estendida: estados mentais estendidos são propriedades de sujeitos híbridos compostos de indivíduos puramente biológicos em conjunto com componentes materiais não biológicos. Esses sujeitos híbridos são numericamente idênticos aos indivíduos puramente biológicos que, em conjunto com características de seus ambientes, passam a constituir os sujeitos híbridos.

Essa proposta de definição da tese da mente estendida captura uma literal extensão da mente de um indivíduo. A entidade que possui mente continua a mesma, mas sua mente passa a ter uma base de superveniência

¹⁶ Inclusive é uma hipótese considerada favoravelmente por Clark e Chalmers (1998, p. 17, tradução nossa): "E o que dizer sobre uma cognição socialmente estendida? Poderiam meus estados mentais serem parcialmente constituídos pelos estados de outros pensadores? Nós não vemos nenhuma razão de porquê não, em princípio. Em um casal extraordinariamente interdependente, é completamente possível que as crenças de um cônjuge desempenhem algum tipo de papel para o outro cônjuge similar àquele desempenhado pelo caderno para Otto. O ponto central é um alto grau de confiança, confiabilidade e acessibilidade."

ampliada. Essa maneira de interpretar a tese reforça sua aproximação com o funcionalismo. O funcionalismo acomoda uma flexibilidade no que diz respeito à composição material da base de superveniência da mente. A tese da mente estendida flexibiliza a extensão (e localização) dessa base: a base de superveniência dos estados mentais pode se estender para além dos limites corporais de uma criatura, incluindo componentes materiais não biológicos (Harris, 2020).

Mas em casos coletivos, como o de Olaf e Olga, já não fica tão claro qual mente biológica é o ponto de partida da extensão. Por que a mente de Olaf deveria ter essa prioridade ontológica? Uma resposta possível seria negar que há uma extensão da mente nos casos coletivos. Talvez há apenas uma interação entre Olaf e Olga. Não há acoplamento e, portanto, não haverá extensão; a mente de Olaf continua restrita aos seus processos biológicos internos e apenas recebe, recorrentemente, testemunhos de Olga. Mas se esse é o caso para Olaf e Olga, por que não seria o caso para Otto e o caderno, você e seu celular, etc.? Por que os critérios de constância, fluidez e endosso deveriam ser interpretados como condições de acoplamento e extensão nesses últimos casos, mas não no primeiro? Por que não é o caso que eles apenas descrevem uma interação distintiva entre uma criatura biológica e recursos não biológicos de seu ambiente semelhante ao que ocorre no caso coletivo?

Ao invés de tomar esse caminho de negação da tese da mente estendida, podemos seguir Harris (2020) e reconsiderar a definição da tese oferecida acima. Harris parte da crítica de Miyazono (2017) que busca, justamente, disputar essa noção de que a mente de um indivíduo possa ser literalmente estendida. Miyazono se inspira no argumento anti-funcionalista de Searle (1980) do “Quarto Chinês”. Nesse caso, imaginamos um indivíduo que está em um quarto com um manual que instrui a produção de certos caracteres em resposta a certos estímulos

(o indivíduo recebe alguns pedaços de papel) que também contêm esses caracteres. Sem saber, o indivíduo que realiza esse processamento estaria estabelecendo uma comunicação em chinês com falantes de chinês localizados fora do quarto. Searle considera bastante implausível que possamos atribuir fluência de chinês para o indivíduo que está dentro do quarto seguindo as ordens do manual sem saber nada de chinês. Entretanto, uma das respostas desenvolvidas pela literatura concede que o indivíduo dentro do quarto não sabe chinês, mas defende que o sistema (indivíduo + manual + quarto) sabe.

Miyazono (2017) argumenta que o mesmo resultado metafísico deveria ser aplicado aos casos de mente estendida. A ideia de que criaturas biológicas individuais podem estender suas mentes, preservando sua identidade numérica ao longo do processo, deveria ser abandonada em favor da consideração de que o sujeito que sustenta os estados mentais e realiza os processos cognitivos é uma nova entidade: o sistema resultante do acoplamento (Otto + caderno; Olaf + Olga). Segundo Harris (2020) as vantagens da resposta do sistema sobre a tese da literal extensão da mente (e identidade numérica da criatura biológica) aparecem na solução do problema do inchaço cognitivo e em considerações sobre persistência.¹⁷

O inchaço cognitivo (Allen-Hermanson, 2013; Rupert, 2009) é a objeção à tese da mente estendida que destaca o excesso de liberalidade nas condições para a extensão da mente. Essa liberalidade acarreta na consideração de casos bastante implausíveis de acoplamento. Harris (2020) cita o exemplo de Giere (2007) sobre o Telescópio Hubble como

¹⁷ De forma bastante similar, o argumento central de Tollefsen (2006) é que o caso de extensão da mente coletiva (com outros indivíduos) consegue acomodar melhor um conjunto de críticas que são feitas à tese da mente estendida, ou seja, a extensão coletiva parece mais plausível que os casos paradigmáticos oferecidos pelas proponentes da tese da mente estendida.

ilustração de um resultado contraintuitivo. Se a tese da mente estendida está correta, então há uma mente que compreenderia o telescópio em órbita, alguns seres humanos e um conjunto de instrumentos localizados na Terra. Novamente, um exemplo que aponta como não fica claro qual mente está sendo estendida. Além de composições muito contraintuitivas, a tese da mente estendida poderia inchar o estoque de crenças de um sujeito de forma inapropriada. Por exemplo, um sujeito que mantém uma relação com uma lista telefônica da forma como Otto se relaciona com seu caderno teria um conjunto implausivelmente grande de crenças sobre números de telefone (Rupert, 2009).

A primeira versão do inchaço cognitivo é respondida pela recusa da literal extensão da mente da criatura biológica envolvida. Justamente pela dificuldade em identificar quem está estendendo sua mente em casos muito complexos de acoplamento, parece mais razoável identificar o surgimento de uma nova entidade, o sistema como um todo, que é o sujeito adequado de atribuição. Da mesma forma, no que diz respeito a segunda versão de inchaço cognitivo, o sistema deve ser o sujeito que mantém o conjunto de crenças e outras propriedades mentais sendo atribuídas: “os sistemas híbridos, compreendendo indivíduos biológicos em conjunto com características de seus ambientes podem ter um número muito mais vasto de crenças do que esses indivíduos biológicos isoladamente” (Harris, 2020, p. 8, tradução nossa).

A resposta do sistema também parece acomodar melhor relações de persistência em casos de acoplamento. No caso de Otto, por exemplo, se tomarmos a literal extensão da mente de Otto com seu caderno, mas mantivermos a identidade de Otto ao longo do processo, isso significaria que Otto se curou do problema de memória. Isso parece pouco razoável. A criatura biológica Otto continua tendo problemas de memória, mesmo quando acoplada a seu caderno, entretanto, essa criatura passa a

integrar um sistema (Otto + caderno) que não possui esse tipo de deficiência. Você continua sem saber qual é o exato número de telefone da sua mãe, mas o sistema composto por você e seu celular pode informar o contato da sua mãe para alguém.

Diante dessas constatações, Harris (2020, p. 10, tradução nossa) apresenta um refinamento na definição da tese da mente estendida:

*Tese da mente estendida**: estados mentais estendidos são propriedades de sujeitos híbridos compostos de indivíduos puramente biológicos em conjunto com outros indivíduos puramente biológicos e/ou componentes materiais não biológicos. Esses sujeitos híbridos não são numericamente idênticos aos indivíduos puramente biológicos que acoplando-se com características de seus ambientes, na forma de outros indivíduos puramente biológicos ou componentes materiais não biológicos, dão origem aos sujeitos híbridos.

Importante salientar que essa formulação continua concedendo que não há uma literal extensão da mente: “(...) nós não podemos nos tornar seres estendidos, mas podemos interagir com seres estendidos. Nós podemos trazer à existência novos seres com mentes cujas mentes sobrevêm de nossos cérebros em conjunto com objetos externos.” (Harris, 2020, p. 8, tradução nossa).

Nesta seção acompanhamos os argumentos de Tollefsen (2006) e Harris (2020) relacionados à aplicação da tese da mente estendida para casos coletivos, evidenciando de que forma a tese da mente estendida é compatível com a tese da mente de grupos. Os autores apresentam a possibilidade de que grupos podem se configurar como sistemas cognitivos que incorporam diferentes seres humanos individuais por meio dos critérios que explicam o acoplamento. No fim de seu artigo, Tollefsen critica explicações reducionistas desse tipo de fenômeno cognitivo social (oferecidas por Wilson [2004]). Entretanto, ela rejeita um modelo emergentista de explicação da relação entre níveis individuais e coletivos,

favorecendo uma explicação em termos de superveniência e realização. Na próxima seção vou apresentar propostas anti-reducionistas emergentistas para a defesa da tese da mente de grupos.

5. Cognição (socialmente) distribuída

A cognição distribuída é outra abordagem que podemos encontrar na literatura que proporciona um caminho de defesa da tese da mente de grupos. As origens de uma parte dessa área de investigação já são marcadamente sociais. Talvez o exemplo mais representativo seja aquele formulado por Edwin Hutchins (1995), um antropólogo cognitivo que, a partir de investigações etnográficas em um navio da marinha americana, constrói um relato bastante detalhado da navegação do navio.

A obra *Cognition in the wild* (Hutchins, 1995) começa com o relato de uma situação de crise. Ao começar a operação para atracar o navio no Porto de San Diego, o engenheiro relata que o tambor de vapor estava perdendo pressão, comprometendo os mecanismos de frenagem da embarcação. A solução era ancorar o navio em mar aberto, para não congestionar o canal de navegação do porto. Mas o navio estava rápido demais para a realização imediata desse procedimento. A fricção da água com o casco do navio seria a única maneira de reduzir a velocidade da embarcação, mas o controle era difícil. O navegador comparou a situação com a perda de motor de uma aeronave ao relatar as dificuldades em conduzir a embarcação sem responsividade no leme. Apesar das dificuldades, a tripulação conseguiu ancorar o navio em lugar apropriado depois de 25 minutos do problema no tambor de vapor e mais de duas milhas percorridas.

Edwin Hutchins (1995, p. 5-6, tradução nossa) destaca o ponto relevante deste relato etnográfico para a tese central do livro:

A ancoragem segura do Palau deveu-se em grande parte à habilidade náutica excepcional da tripulação, especialmente do navegador. Mas nenhum indivíduo agindo sozinho na ponte de comando, nem o capitão, nem o navegador, nem o contramestre que supervisiona a equipe de navegação - poderia ter mantido o controle do navio e conduzi-lo com segurança para a ancoragem. Muitos tipos de pensamento foram necessários para executar esta tarefa. Alguns deles aconteciam em paralelo, alguns em coordenação com outros, alguns dentro das cabeças dos indivíduos e alguns claramente dentro e fora das cabeças dos participantes.

A partir desse exemplo, pode-se notar que, em grande medida, a noção de cognição distribuída já incorpora elementos sociais e levanta a questão de se algumas propriedades, atividades, processos e estados cognitivos podem ser atribuídos ao grupo ou sistema sem que sejam possuídos por nenhuma de suas partes.

Georg Theiner (2010, 2017) desenvolve uma argumentação a favor da tese da mente de grupos precisamente por meio da identificação de propriedades emergentes de sistemas com cognição socialmente distribuída. Uma parte significativa da proposta envolve uma discussão conceitual de *emergência*, visto seu usual descrédito em virtude de um aparente compromisso ontológico com propriedades misteriosas e não explicáveis cientificamente.¹⁸ Um elemento central da reabilitação do conceito de emergência se dá na tentativa de explicar o slogan “é mais que a soma das partes”, ou seja, fornecer uma distinção entre sistemas (entidades que têm propriedades emergentes) e meros agregados (entidades que não as têm).

¹⁸ Theiner (2014) cita que um movimento de defesa da tese da mente de grupos que surgiu na passagem do século XIX para o XX tentava explicar as características dos grupos que não podiam ser reduzidas a agentes individuais por meio de um vocabulário emergentista vitalista. A adoção dessa estratégia foi bastante prejudicial para a tese da mente de grupos: “os vitalistas acreditavam que a vida é produto de uma força orgânica misteriosa (*vis vitalis*) que é fundamentalmente diferente dos princípios físico-químicos que governam as coisas inanimadas. Por conta dessa estreita associação, o fim do vitalismo como resultado da síntese evolucionária moderna na biologia significou que os conceitos de mente de grupo ou agência de grupo fossem igualmente banidos do domínio de discurso cientificamente respeitável” (Theiner, 2014, p. 301, tradução nossa).

Theiner (2010) busca estabelecer essa distinção com base em Wimsatt (1986). Os critérios de Wimsatt capturam a noção geral de emergência (o resultado conjunto de duas ou mais partes de um sistema difere do resultado agregado dessas partes operando independentemente) e de sinergia (o comportamento de partes que tenham interação mútua é diferente do comportamento dessas partes em isolamento) (Theiner, 2017). Esses critérios são construídos a partir de uma perspectiva negativa, a partir de falhas de agregação que revelariam algum tipo de dependência que propriedades emergentes teriam com relação à organização do sistema. Um sistema seria um mero agregado caso não tivesse propriedades emergentes, ou seja, caso suas propriedades se alterassem exclusivamente devido à (i) substituição de suas partes; (ii) adição ou subtração de partes (que geram apenas alterações quantitativas); (iii) reorganização de suas partes; e (iv) inexistência de interações cooperativas ou inibitórias entre as partes (Theiner, 2017).

Podemos sintetizar a ideia central deste modo de caracterizar sistemas com propriedades emergentes em virtude de falhas de agregação como destacando o fato de que o comportamento do sistema não poderá ser entendido por uma análise reducionista que apenas identificaria as atividades das unidades componentes operacionais de maneira independente. Assim, uma estratégia simples de decomposição funcional não será um modelo explicativo adequado para o comportamento de um sistema que possui propriedades emergentes, isto é, cujos comportamentos são resultados da interação de suas unidades componentes operacionais. Ainda que essas estratégias reducionistas sejam uma ferramenta heurística relevante para a produção de explicações simples, elas podem acarretar em falácias como “a mente não é nada além da atividade neural” ou “comportamento social não é nada além da ação de indivíduos” (Wimsatt, 1997).

Um dos exemplos de sistemas com cognição socialmente distribuída apresentado por Theiner (2010) é o fenômeno da memória transativa estudada pela psicologia social (Wegner, 1985; 1987¹⁹). Em alguns casos, certos grupos (famílias, equipes de trabalho) desenvolvem uma divisão de trabalho cognitivo para resgatar informações, para se lembrar conjuntamente. As tarefas de alocação, codificação, recuperação, elaboração e compartilhamento de informações, características da memória, são realizadas conjuntamente, fora das cabeças dos indivíduos envolvidos.

O elemento fundamental de uma memória transativa é a divisão do trabalho cognitivo que evidencia os aspectos coletivos envolvidos. Os processos de codificação e alocação de informações levam em conta as especialidades dos membros do grupo. Essas atividades exigem que os membros tenham conhecimento de especificidades e especialidades de cada indivíduo pertencente ao grupo, promovendo uma distribuição do trabalho cognitivo que alivia as responsabilidades individuais. A identificação de especialistas também é relevante na operação de resgate das informações. Quando uma informação for necessária, espera-se que os membros saibam quem detém esse tipo de informação.²⁰

Tollefsen (2006, p. 145) ilustra o processo, destacando suas características coletivas, a partir do exemplo de Olaf e Olga:

¹⁹ Essa também foi a fonte central para corroborar o caso Olaf e Olga em Tollefsen (2006), discutido na seção anterior.

²⁰ Liang, Moreland e Argote (1995) apresentam sistemas de memórias transativas de um modo um pouco mais formal, mas que pode esclarecer um pouco mais as atividades características desse tipo de sistema. Um sistema de memória transativa seria uma propriedade de segunda-ordem ancorada por três processos de primeira-ordem, também coletivos: "Diferenciação de memória (M1), a tendência dos membros do grupo em se especializarem na recordação do processo coletivo; credibilidade de tarefa (M2), quanto um membro confia na expertise de outro (associada a comportamentos tais como menor necessidade de afirmar expertise, melhor aceitação de sugestões procedimentais, menos críticas); e coordenação de tarefa (M3), a habilidade dos membros do grupo em trabalharem juntos coordenadamente (mensurada, p.e., pela menor necessidade de planejamento explícito, menos mal entendidos, maior cooperação)." (Theiner, 2010, p. 93, tradução nossa)

[I]magine que o resgate de informações e sua reorganização é feito por [Olga] e Olaf conjuntamente. Olaf não se lembra do sobrenome; [Olga] não se lembra do primeiro nome. Olaf se lembra onde eles se encontraram. [Olga] se lembra do porquê eles se encontraram. Através de um processo de deliberação conjunta eles resgatam conjuntamente, por meio de um processo de reconstrução conjunta, o nome da pessoa. O processo de resgate, então, é ativo e não pode ser encontrado "dentro" da cabeça de [Olga] ou Olaf. De fato, ele é feito por meio de uma atividade conjunta de discussão e deliberação que ocorre entre [Olga] e Olaf.

Cabe destacar que Theiner (2010) identifica que sistemas de memória transativas são exemplos de falhas de agregação, isto é, são exemplos de processos que instanciam sistemas que possuem propriedades emergentes. Esse tipo de sistema não acomoda troca entre as partes, visto que há uma especialização na alocação de memórias. Portanto, se alguém com uma informação crucial for retirado ou substituído (critérios (i) e (ii)), o funcionamento do grupo será afetado. Além disso, experimentos indicaram que grupos que desenvolvem sistemas de memória transativas são mais produtivos que grupos de pessoas treinadas separadamente e depois reunidas (Lewis et. al. 2005; 2007). Essa evidência aponta para uma negação de que os agregados não diferem quando suas partes são reorganizadas (critério (iii)). Finalmente, como cada membro tem conhecimento de especificidades e especialidades de cada indivíduo do grupo, seus processos individuais de aquisição, resgate e comunicação de memórias é afetado por fazer parte de um sistema de memórias transativas. Portanto, há interações cooperativas e inibitórias entre as partes componentes deste tipo de sistema (critério (iv)). Desta forma, nenhum dos critérios de agregação de Wimsatt são satisfeitos nos casos de sistemas de memórias transativas, evidenciando que se trata de um processo com resultados emergentes.

Outra obra que apresenta uma defesa da tese da mente de grupos por meio de aspectos distribuídos da cognição é *Macro cognition* de Bryce

Hubner (2014). A base do argumento de Hubner é uma analogia entre sistemas de cognição distribuída: entre o funcionamento do sistema de cognição humano e o sistema de cognição de grupo. Hubner parte de uma noção de cognição distribuída um pouco diferente daquela preconizada por Hutchins. A noção de cognição distribuída nas Ciências Cognitivas se caracteriza pela adoção de modelos explicativos para a cognição humana orientados por uma arquitetura de um sistema composta por muitas unidades computacionais de processamento de informação altamente especializadas (destinadas a uma tarefa muito específica, como, por exemplo, detectar sombras nas representações visuais, detectar formas nas representações visuais, etc.). Essas unidades altamente especializadas são integradas e coordenadas de forma a produzir comportamentos flexíveis, habilidosos e orientados a objetivos por parte do sistema como um todo. Hubner (2014, p. 63, tradução nossa) sintetiza esses pressupostos das Ciências Cognitivas da abordagem conexionista da seguinte forma:

Para explicar a operação de uma mente individual, devemos adotar uma perspectiva macrocognitiva. Devemos postular vários sistemas computacionais, constatar que eles desempenham uma enorme variedade de tarefas computacionais, reparar que eles frequentemente fazem isso independentemente uns dos outros, avaliar que eles frequentemente funcionam paralelamente e perceber que eles frequentemente operam baseados em estruturas altamente localizadas de interfaces entre sistemas computacionais subpessoais.

Note que não há um elemento propriamente social nessa noção de cognição distribuída. Similar à proposta de Theiner, há um sistema composto de unidades que apresenta certas propriedades emergentes, mas essas unidades não são indivíduos, mas sim unidades computacionais de processamento de informação altamente especializadas. A proposta dessa abordagem em Ciências Cognitivas visa

aproximar o processamento cognitivo do cérebro ao processamento paralelo de computadores. O aspecto mais importante dessa proposta é a decomposição funcional, ou seja, a concepção de que a mente humana é uma espécie de sociedade constituída por muitos agentes organizados em uma estrutura. A ideia é que a mente seria formada por muitos agentes menores, as unidades computacionais de processamento de informação especializadas, sendo que cada um deles não teria uma mente.

A estratégia geral de Hubner é uma analogia entre dois tipos de sistemas de cognição distribuída, defendendo a hipótese de que podem haver sistemas de diferentes níveis. Há o sistema da mente humana, descrito pelas Ciências Cognitivas e há o sistema da mente de grupo, semelhante aos casos de cognição socialmente distribuída descritos por Hutchins e Theiner. Para ilustrar melhor essa estratégia argumentativa, Hubner faz referência a caracterização do sistema cognitivo humano proposta por Marvin Minsky (1988), onde:

(...) explicações psicológicas fornecem uma estrutura hierárquica de decomposições mecanicistas, oferecendo explicações psicológicas efetivas e ontologicamente inocentes ao postular sucessivos níveis de homúnculos mais "estúpidos" até que a hierarquia inevitavelmente chegue ao fundo, em homúnculos tão estúpidos que eles poderiam ser implementados por botões de liga/desliga (ou seus equivalentes). O funcionalista homuncular trata mentes individuais como "sociedades" de agentes, organizados em hierarquias corporativas, onde cada homúnculo desempenha apenas um conjunto limitado de tarefas computacionais (Hubner, 2014, p. 64-65, tradução nossa).

Hubner propõe inverter a lógica dessa alegoria. Ao invés de uma sociedade de agentes originar uma mente individual, ele tenta defender que um conjunto de indivíduos (com mentes) pode conferir a uma hierarquia corporativa uma mente. Assim, adotando as estratégias explicativas das teorias das Ciências Cognitivas, Hubner pretende abrir

a possibilidade para identificar quando estados e processos mentais coletivos emergem: quando são o resultado da interação de indivíduos incorporados em arquiteturas computacionais compostas por agentes humanos e ferramentas para comunicação.

Para esclarecer a importância da decomposição funcional no argumento geral de Hubner, precisamos voltar aos modelos de explicação funcionalista, como aqueles apresentados na seção 3. Novamente, vamos partir da aplicação de modelos de explicação funcionalista para casos mais simples do que a explicação da mente. Para entendermos um sistema como um motor de combustão, devemos identificar seus componentes. Um deles é o carburador, um componente responsável pelo controle da mistura de combustível e ar dentro do motor. Sem uma teoria prévia do que é um carburador, uma noção dessa função desempenhada por um carburador em um motor, teríamos muita dificuldade em identificar um carburador em um motor específico. Essa teoria prévia (i.e. perfil funcional) sobre carburadores deveria ser suficientemente estável e projetável para que pudéssemos identificar carburadores em outros modelos de motores de combustão que, possivelmente, poderiam ser construídos de várias maneiras distintas.

Hubner acredita que a explicação da mente procede de modo análogo a esse exemplo sobre motores de combustão. Psicólogos e cientistas cognitivos estão interessados em um tipo de sistema específico, tipicamente um indivíduo humano. Mais precisamente, esses pesquisadores estão interessados em descobrir os mecanismos (componentes) que explicam os padrões de comportamento desse tipo de sistema específico. Para isso, eles precisam de teorias sobre os componentes do sistema, como crenças, desejos e intenções, para identificar onde estão esses componentes dentro do sistema.

Aqui, Hubner segue a tese da postura intencional de Daniel Dennett (1987). A postura intencional é a maneira como procedemos para explicar e prever o comportamento de uma entidade que consideramos ser um agente racional. Ela se assemelha a nossa prática cotidiana de atribuição de crenças, desejos e intenções a outros seres humanos à nossa volta. Seres humanos são exemplos paradigmáticos de agentes racionais. Eles agem racionalmente em virtude de seguirem um padrão normativo de razões. Esse padrão normativo é uma espécie de ponto de vista racional constituído de crenças, desejos, intenções, etc., que nos permite identificar quais são as razões pelas quais o indivíduo age da maneira como age. Somos relativamente bem-sucedidos nessa tarefa, pois também somos agentes racionais, ou seja, compartilhamos com outros agentes racionais um padrão normativo que guia nossos comportamentos.

Utilizando a postura intencional e, portanto, caracterizando o que são crenças, desejos e intenções (por meio de seus perfis funcionais, pelo modo como esses elementos guiam o comportamento do sistema), os cientistas cognitivos passam a descobrir os mecanismos que produzem os comportamentos flexíveis, habilidosos e orientados a objetivos característicos de um sistema como um indivíduo humano, i.e. realizam a decomposição funcional desse tipo de componente do sistema. A investigação sobre esses mecanismos (seja o carburador em um motor de combustão, sejam crenças e desejos em um indivíduo humano) pode explicar como um sistema faz aquilo que ele faz. Adotando a perspectiva conexionista das Ciências Cognitivas, os componentes de um sistema como um indivíduo humano são mecanismos subpessoais que operam processos computacionais e algorítmicos organizados em uma arquitetura característica. Esses mecanismos subpessoais computacionais revelam o que está por trás das nossas atribuições de crenças, desejos e intenções, das nossas atribuições baseadas na postura

intencional, identificando quais são os realizadores físicos desses estados mentais que explicam, a partir de uma perspectiva mecanicista causal, o comportamento padrão do sistema sob investigação.

Essa descrição do modelo de explicação da Psicologia e das Ciências Cognitivas pode dar a entender que a proposta de Hubner é reducionista, privilegiando a identificação das estruturas internas causais na produção (e explicação) de comportamentos flexíveis, habilidosos e orientados a objetivos, isto é, de comportamentos inteligentes de seres humanos. Hubner reconhece o papel de destaque que esses mecanismos têm em uma explicação naturalista, ancorada em pressupostos das Ciências Cognitivas. A ausência de uma descrição dos componentes do sistema impede que entendamos como ele se comporta e não teremos elementos para defender que esse sistema instancia processos cognitivos e, conseqüentemente, que ele tenha uma mente.

Entretanto, esses pressupostos não eliminam a possibilidade de que a proposta de Hubner incorpore propriedades emergentes. Sistemas cognitivos manifestam propriedades que não são instanciadas por seus componentes e que emergem por conta da forma como esses mecanismos estão organizados no interior do sistema. As propriedades emergentes são relevantes na medida em que explicam determinados comportamentos do sistema que não podem ser apreendidos pela análise isolada da operação dos componentes do sistema, por uma mera decomposição funcional que respeita os critérios de Wimsatt, por exemplo (Bechtel; Richardson, 1993; Bechtel; Abrahamsen, 2002). Desta forma, essa proposta de explicação para fenômenos psicológicos envolve uma espécie de equilíbrio reflexivo, onde se faz necessário um movimento de vai e vem entre a articulação das funções que os componentes do sistema têm (i.e. seus perfis funcionais) e as explicações mecanicistas-causais da operação desses componentes.

Assim, há uma importância não negligenciável para o uso da postura intencional (nossas teorias sobre crenças, desejos, intenções, etc.) no desenvolvimento de uma explicação adequada do comportamento de um sistema cognitivo (como também havia uma importância para a adoção prévia de uma teoria sobre carburadores para compreender um motor de combustão).

Hubner defende esse modelo explicativo a partir de um argumento de Dennett (1987), onde:

(...) até mesmo um marciano laplaciano, que pode prever o movimento de todas as partículas físicas do universo, não perceberia padrões perfeitamente objetivos se não adotasse a postura intencional como base para explicações psicológicas. [Dennett] afirma que tais marcianos não seriam capazes de ver o número indefinido de padrões únicos de movimentos físicos que poderiam ser substituídos sem a perturbação do comportamento direcionado a um objetivo. (...) Ao nível físico, esses estados são heterogêneos, então eles não poderiam ser classificados como pertencendo ao mesmo tipo particular, exceto por meio de uma explicação psicológica. (Hubner, 2014, p. 94, tradução nossa)

Um papel determinante que a postura intencional desempenha no desenvolvimento de explicações psicológicas é a acomodação da múltipla realizabilidade de estados mentais. A definição do perfil funcional desses estados, a descrição do papel que eles desempenham no sistema, é importante para a localização dos seus realizadores, dos mecanismos físicos específicos que são ativados em uma instanciação particular desses estados, que podem ser completamente diferentes em entidades diferentes. Hubner (2014, p. 64, tradução nossa) nos apresenta outra forma de compreender a relação entre esses dois níveis de explicação necessários para a caracterização de um sistema cognitivo:

Ao adotar a postura intencional, nós tratamos estados mentais como objetos abstratos que são individuados por suas propriedades intencionais e funcionais. Postular tal *abstracta* nos permite decompor as capacidades

cognitivas em complexos de mecanismos intencionalmente especificados, os quais podem explicar a presença de uma capacidade comportamental ou cognitiva.

É essa noção geral de que a mente é caracterizada por uma arquitetura de um sistema composta por muitas unidades computacionais de processamento de informação altamente especializadas capaz de produzir um comportamento flexíveis, habilidosos e orientados a objetivos que abre a possibilidade para a existência de mentes de grupos. Grupos poderão se enquadrar enquanto sistemas cognitivos se também apresentarem essa arquitetura, isto é, se eles exibirem uma variedade de comportamentos que sejam apropriadamente caracterizados como intencionais e que sejam gerados por um sistema computacional integrado.

Essa arquitetura pode ser distinta daquela presente em seres humanos individuais, isto é, ela não precisa ser, necessariamente, composta de neurônios e outras entidades hidrocarbonadas características de um sistema cognitivo orgânico. Para ser um sistema cognitivo, basta que um grupo exiba um comportamento coletivo complexo que seja explicado a partir de propriedades emergentes de um sistema computacional integrado. Muito possivelmente, a descrição desse tipo de arquitetura dará ênfase a seres humanos individuais e aparatos tecnológicos como seus componentes, esses serão os realizadores, os mecanismos físicos específicos ativados na instanciação de estados cognitivos de grupos. Sintetizando, Hubner (2014, p. 96, tradução nossa) destaca os seguintes itens para a identificação de uma mente em um grupo:

- 1) Especificação dos tipos de informações aos quais uma coletividade é sensível;

- 2) Decomposição funcional dos estados cognitivos que a coletividade é capaz de ter e dos processos cognitivos que ela é capaz de executar; e
- 3) Descrição dos procedimentos computacionais que implementam esses estados e processos mentais.

É importante destacar que a proposta de Hubner segue parte do caminho funcionalista baseado na postura intencional, de forma muito similar às propostas discutidas na seção 3, ou seja, parte da identificação de que o comportamento de (alguns) grupos exibe o nível de racionalidade característico de um agente. Entretanto, Hubner considera as propostas interpretativistas insuficientes. Não podemos ficar restritos à postura intencional, é fundamental abrir a caixa preta e identificar quais são os mecanismos que operam os processamentos computacionais necessários para a atribuição de mente a um sistema, isto é, os mecanismos causais que explicam a produção de comportamentos flexíveis e habilidosos por parte de um grupo (ver também Strohmaier [2020]).

Conclusão

Neste texto apresentei uma visão geral de propostas que podem ser utilizadas na defesa da tese da mente de grupos. O objetivo central foi bastante modesto, apenas mostrar que a tese da mente de grupos é compatível com muitas teorias contemporâneas sobre a natureza da mente sendo discutidas na Filosofia. Importante notar que o recorte da discussão aqui apresentado é bastante positivo. A ausência de posições críticas evidencia que essa apresentação é bastante parcial (talvez a posição crítica mais direta seja Rupert [2011, 2014, 2019]). Acredito que, de modo geral, alguém contrário à tese toma sua negação como pressuposto plausível, julgando desnecessária uma argumentação contrária e por isso não há tantas obras focadas na refutação da tese).

Destaco que algumas respostas aos argumentos Fisicalista e da Consciência, apresentados na seção 2, aparecem ao longo do texto. O elemento central de defesa contra esses dois ataques está ancorado na adoção de uma posição funcionalista sobre a natureza da mente. De modo muito direto, o funcionalismo rejeita o pressuposto central do argumento Fisicalista (“A mente depende do cérebro”) que preserva a intuição de que a mente estaria necessariamente relacionada com algum tipo específico de constituição material. De modo um pouco mais indireto, o funcionalismo também levanta dúvidas sobre o pressuposto central do argumento da Consciência (“A mente depende da consciência”). Em grande medida por conta do funcionalismo, a Filosofia da Mente começa a considerar a distinção entre consciência de acesso e consciência fenomênica (Block, 1995). A segunda é nossa noção intuitiva de consciência, caracterizada pela rica experiência qualitativa tão distintiva do contato que temos com nossas próprias mentes. A consciência de acesso contemplaria a disponibilidade de certos estados em guiar a fala e a ação, caracterizada, portanto, por um perfil funcional. Com certeza, o argumento da Consciência diz respeito à consciência fenomênica, entretanto, a possibilidade de que exista outro tipo de consciência é o que traz alguma dúvida sobre se o grau de restrição imposto pelo pressuposto de que “A mente depende da consciência [fenomênica]” é adequado.

Além do funcionalismo, outra característica muito saliente das defesas da tese da mente de grupo é a identificação de propriedades próprias do nível coletivo (sejam elas emergentes, supervenientes ou realizadas, a relação metafísica entre os níveis não é consensual na literatura). A identificação desse tipo de propriedade se configura como uma das principais evidências a favor de uma posição realista sobre mente de grupos (ou, ao menos, uma legitimidade para o campo de

investigação, seguindo o argumento a favor das ciências especiais de Fodor [1974]). Apesar de recorrente, nem todas as estratégias abraçam essa visão estratificada da realidade. Um exemplo de alternativa não-emergentista, que não pôde ser explorado aqui, é o pampsiquismo. Essa consideração é relevante, pois evidencia que as diferentes defesas da tese da mente de grupos podem não ser compatíveis entre si e que diferentes teorias sobre a natureza da mente podem acarretar em diferentes teorias sobre a realidade social.²¹

Referências

- ALLEN-HERMANSON, S. Superdupersizing the Mind: Extended Cognition and the Persistence of Cognitive Bloat. *Philosophical Studies*, v.164, n.3, p. 791-806, 2013.
- BECHTEL, W.; RICHARDSON, R. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton: Princeton University Press, 1993.
- BECHTEL, W.; ABRAHAMSEN, A. *Connectionism and the Mind: Parallel Processing, Dynamics, and Evolution in Networks*. Oxford: Basil Blackwell, 2002.
- BLOCK, N. On a Confusion About a Function of Consciousness. *Behavioral and Brain Sciences*, v.18, p. 227-272, 1995.
- BRICKER, P. Ontological Commitment. In: ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*, 2014.
- CARVALHO, E. A Tese da Mente Estendida à Luz do Externismo Ativo: Como Tornar Otto Responsivo a Razões? *Trans/Form/Ação*, v.43, n.3, p. 143-166, 2020.

²¹ Gostaria de agradecer ao Anderson Fonseca pelo convite em participar da obra, inclusive das atividades de organização do volume, fico muito contente de compartilhar o feito de trazer a público essa relevante produção.

Também agradeço à Adriana Oliveira Alves pela revisão da versão preliminar deste texto. Suas sugestões foram muito importantes para tornar esse difícil texto um pouco mais compreensível.

- CICHOSKI, L.; RUIVO, J. Agregação de Juízo na Epistemologia Social: a Proposta de Christian List e Philip Pettit. In: CARVALHO, M. et al. (orgs.). *Pragmatismo, Filosofia Analítica e Filosofia da Mente*. São Paulo: ANPOF, p. 273-302, 2015.
- CICHOSKI, L.; RUIVO, J. Epistemologia Coletiva: Crença, Justificação e Conhecimento de Grupo. *Veritas*, v. 62, n. 3, p. 508-539, 2017.
- DAHAN, O. The Problem of Other (Group) Minds. *Philosophia*, v.45, n.3, p. 1099-1112, 2017.
- DENNETT, D. *The Intentional Stance*. Cambridge: MIT press, 1987.
- FRENCH, P. The Corporation as a Moral Person. *American Philosophical Quarterly*, v.16, n.3, p. 207-215, 1979.
- FRENCH, P. *Collective and Corporate Responsibility*, New York: Columbia University Press, 1984.
- FRENCH, P. *Corporate Ethics*. Orlando: Harcourt Brace, 1995.
- FODOR, J. Special Sciences, or Disunity of Science as a Working Hypothesis. *Synthese*, v.28, n.2, p. 97-115, 1974.
- FONSECA, A. *A Colônia de Formiga Está Consciente? Uma análise sob a ótica do pampsiquismo*. 115f. Dissertação (Mestrado em Filosofia), Universidade Federal de São João del-Rei, 2022.
- FONSECA, A. É a Colônia de Formigas um Organismo Consciente? *Griot*, v.23, n.1, p.70-86, 2023.
- GILBERT, M. Modelling Collective Belief. *Synthese*, v. 73, n.1, p. 185-204, 1987.
- GILBERT, M. Belief and Acceptance as Features of Groups. *ProtoSociology*, v. 16, p. 35-69, 2002.
- GOUVEA, R. Physicalism Without Identity. *Trans/form/ação*, v.43, n.2, p. 253-280, 2020.
- GIERE, R. Distributed Cognition Without Distributed Knowing. *Social Epistemology*, v. 21, n.3, p. 313-320, 2007.

- HARRIS, R. Group Minds as Extended Minds. *Philosophical Explorations*, v.23, n.3, p. 1-17, 2020.
- HUBNER, B. *Macrocognition*. Oxford: Oxford University Press, 2014.
- HUTCHINS, E. *Cognition in the Wild*. Cambridge: MIT Press, 1995.
- JACOB, P. Intentionality. In: ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*, 2023.
- KRAMER, M. What it Might be Like to be a Group Agent. *Neuroethics*, v.14, n.3, p. 437-447, 2021.
- KORHAUSER, L.; SAGER, G. Unpacking the Court. *Yale Law Journal*, v.96, p. 82-117, 1986.
- LEWIS, K.; LANGE, D.; GILLIS, L. Transactive Memory Systems, Learning, and Learning Transfer. *Organization Science*, v.16, n.6, p. 581-598, 2005
- LEWIS, K.; BELLIVEAU, M.; HERDON, B.; KELLER, J. Group Cognition, Membership Change, and Performance: Investigating the Benefits and Detriments of Collective Knowledge. *Organizational Behavior and Human Decision Processes*, v. 103, p. 159-178, 2007.
- LIANG, D.; MORELAND, R.; ARGOTE, L. Group Versus Individual Training and Group Performance: The Mediating Role of Transactive Memory. *Personality and Social Psychology Bulletin*, v. 21, p. 384-393, 1995.
- LINDEMAN, D. Propositional Attitudes. *The Internet Encyclopedia of Philosophy*, 2023. Disponível em: <<https://iep.utm.edu/prop-ati/>>. Acesso em: 20 de dez. 2023
- LIST, C; PETTIT, P. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press, 2011.
- LIST, C. What is it Like to be a Group Agent? *Noûs*, p. 295-319, 2016.
- LIST, C. Social Choice Theory. In: ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*, 2022.
- MAY, L.; HOFFMAN, S. (eds.) *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics*. Savage: Rowman and Littlefield, 1991.

- MERLEAU-PONTY, M. *Fenomenologia da Percepção*. São Paulo: Martins Fontes, 1999.
- MINSKY, M. *Society of the Mind*. New York: Simon and Schuster, 1988.
- MIYAZONO, K. Does Functionalism Entail Extended Mind? *Synthese*, v. 194, n.9, p. 3523-3541, 2017.
- NAGEL, T. Como é ser um Morcego? *Cadernos de História e Filosofia da Ciência*, v. 15, n. 1, p. 245-262, 2005.
- QUINTON, A. Social Objects. *Proceedings of the Aristotelian Society*, v.75, p. 1- 27, 1975.
- RAMSEY, W. Eliminative Materialism. In: ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*, 2019.
- ROELOFS, L. *Combining Minds: How to Think About Composite Subjectivity*. Oxford: Oxford University Press, 2019.
- RUIVO, J. *Crença de grupo: uma Introdução à Epistemologia Coletiva*. 102 f. Tese (Doutorado em Filosofia) – PUCRS, Porto Alegre, 2017.
- RUPERT, R. *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press, 2009.
- RUPERT, R. Empirical Arguments for Group Minds: A Critical Appraisal. *Philosophy Compass*, v. 6, n. 9, p. 630-639, 2011.
- RUPERT, R. Against Group Cognitive States. In: PREYER, G.; HINDRICKS, F.; CHANT, S. (eds.) *From Individual to Collective Intentionality: New Essays*. Oxford: Oxford University Press. p. 97-111, 2014.
- RUPERT, R. Group Minds and Natural Kinds. *Avant*, v. 10, n. 3, p. 1-28, 2019.
- PETTIT, P. Groups with Minds of Their Own. In: SCHMITT, F. (ed.) *Socializing Metaphysics*. Lanham: Rowman & Littlefield. p. 167-193, 2003.
- POLGER, T. Functionalism. *The Internet Encyclopedia of Philosophy*, 2023. Disponível em: < <https://iep.utm.edu/functionism>>. Acesso em: 15 de jun. 2023.

- SCHWITZGEBEL, E. If Materialism is True, the United States is Probably Conscious. *Philosophical Studies*, v. 172, n.7, p. 1697-1721, 2015
- SCHLOSSER, M. Agency. In: ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*, 2019.
- SEARLE, J. Minds, Brains, and Programs. *Behavioral and Brain Sciences*, v. 3, n. 3, p. 417-457, 1980.
- SMILEY, M. Collective Responsibility. In: ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*, 2022.
- STROHMAIER, D. Two Theories of Group Agency. *Philosophical Studies*, v. 177, n. 7, p. 1901-1918, 2020.
- SZANTO, T. How to Share a Mind: Reconsidering the Group Mind Thesis. *Phenomenology and the Cognitive Sciences*, v.13, n.1, p. 99-120, 2014.
- TALBERT, M. Moral Responsibility. In: ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*, 2024.
- THEINER, G.; O'CONNOR, T. The Emergence of Group Cognition. In: CORRADINI, A.; O'CONNOR, T. (eds), *Emergence in Science and Philosophy*, New York: Routledge. p. 78-117, 2010.
- THEINER, G. A Beginner's Guide to Group Minds. In: SPREVAK, M.; KALLESTRUP, J. (eds.) *New Waves in Philosophy of Mind*. Palgrave-macmillan. p. 301-22, 2014.
- THEINER, G. Groups as Distributed Cognitive Systems. In: Groups as Distributed Cognitive Systems. In: JANKOVIC, M.; LUDWIG, K. (eds.) *The Routledge Handbook of Collective Intentionality*. New York: Routledge, p. 233-248, 2017.
- TOLLEFSEN, D. Challenging Epistemic Individualism. *ProtoSociology*, v. 16, p.86-117, 2002.
- TOLLEFSEN, D. From Extended Mind to Collective Mind. *Cognitive Systems Research*, v.7, n.2, p. 140-150, 2006.
- TOLLEFSEN, D. *Groups as Agents*. Cambridge: Polity Press, 2015.

WEGNER, D.; GIULIANO, T.; HERTEL, P. Cognitive Interdependence in Close Relationships. In: ICKES, W. (ed.) *Compatible and Incompatible Relationships*. New York: Springer, p. 253-276, 1985.

WEGNER, D. Transactive Memory: A Contemporary Analysis of the Group Mind. In: MULLER, B.; GOETHALS, G. (eds.) *Theories of Group Behavior*. New York: Springer, p. 185-208, 1987.

WILSON, R. *Boundaries of the Mind*. Cambridge: Cambridge University Press, 2004.

WIMSATT, W. Forms of Aggregativity. In: GRENE, M.; DONAGAN, A.; PEROVICH, A.; WEDIN, M. (eds) *Human Nature and Natural Knowledge*. Dordrecht : Reidel, 1986.

8

MENTES DE GRUPOS E TIPOS NATURAIS ^{1 2 3}

Robert Rupert

Tradução: Anderson Fonseca, Jeferson

Diello Huffermann, Luiz Paulo Cichoski

1. Introdução

Este ensaio aborda a questão da possibilidade de que grupos de humanos individuais - por exemplo, tribunais, corporações e sindicatos - tenham estados mentais próprios, para além dos estados mentais de seus membros. Minha orientação metodológica será naturalista e, portanto, eu tratarei o assunto como uma questão de Filosofia da Ciência, especialmente dentro da Filosofia das Ciências Cognitivas.⁴

¹ Tradução do texto: RUPERT, Robert. Group minds and Natural Kinds. *AVANT*, Bydgoszcz, v. X, n. 3, p. 1-28, 2019; DOI: 10.26913/avant.2019.03.08. Agradecemos ao autor e aos editores da revista pela autorização da publicação desta tradução. (N.T.)

² Essa publicação da revista *AVANT* foi financiada pelo Ministério de Ciência e Educação Superior (da Polônia) a partir dos recursos para a disseminação de pesquisa (DUN) no âmbito de atividades de publicação, contrato nº 711/P-DUN/2019, período de implementação: 2019-2020.

³ Este artigo é derivado de um manuscrito que previamente circulou nomeado "Indivíduos como mentes de grupos, mentes de grupos como indivíduos". Porções significativas desse material foram apresentadas na Universidade Macquarie, Universidade de Auckland, Universidade de Edimburgo, Universidade Bochum Ruhr, Collective Intentionality IX na Universidade de Indiana, *Avant: Trends in Interdisciplinary Studies III* em Lublin, Universidade Johns Hopkins, Universidade Humboldt e no Grupo de Discussão em Filosofia da Mente da Universidade de Nova York. Muito obrigado aos vários membros da audiência e participantes de grupos por seus comentários. Agradecimentos especiais a Georg Theiner.

⁴ Uma vasta literatura sobre mente de grupo foi conduzida predominantemente por reflexões sobre formas ordinárias de atribuição de estados mentais. Por exemplo, ao descrever seus pressupostos metodológicos, Margaret Gilbert (2003) diz, "É suficiente dizer que minhas próprias tendências seguem o discurso cotidiano, da forma como eu o entendo. Se tal-e-tal fenômeno é referido como sendo, séria e literalmente, uma crença, é difícil argumentar que ele, afinal de contas, não seja uma crença como ordinariamente compreendida." (p. 103). Eu vou deixar essa literatura de lado. Eu gostaria de saber se grupos realmente têm estados cognitivos próprios e, para minha mente, isso requer a compreensão do trabalho científico sobre cognição, em grande medida desconsiderando aquilo que o leigo tende a pensar sobre questões relativas ao mental e ao cognitivo (reconhecendo, no entanto, que os padrões de uso ordinários podem, no fim das contas, ajudar a determinar, não exatamente onde as juntas da natureza estão, mas como rotular as categorias que forem obtidas - veja Stich, 1996, cap. 1). Por razões similares, eu deixo de lado investigações empíricas sobre as tendências ordinárias na atribuição de

Desta forma, eu me concentrarei em questões como “Em virtude do que um estado ou processo é cognitivo?” e “Quais tipos de evidência sustentam a afirmação de que grupos têm estados cognitivos ou envolvem-se em processamento cognitivo?” A respeito da última questão, eu terei interesse particular em casos nos quais o processamento do grupo difere significativamente do processamento de humanos individuais, perguntando qual igualdade (*sameness*) de tipo - ou, alternativamente, quais tipos de evidências sustentam a afirmação de igualdade de tipo - diante de tais diferenças.

Certas correntes das Ciências Cognitivas e das Ciências Sociais enquadram a discussão. Ao longo das últimas três décadas, ficou cada vez mais claro que coleções de processos distribuídos produzem uma porção substancial dos comportamentos humanos inteligentes (Dennett, 1991; Brooks, 1999; Gendler, 2008; Evans; Frankish, 2009; Clark, 1997). Durante aproximadamente o mesmo período, um número crescente de pesquisadores trabalhando nas Ciências Cognitivas e nas Ciências Sociais, assim como na Filosofia da Mente empiricamente orientada, têm argumentado que alguns grupos de humanos individuais têm estados mentais ou cognitivos próprios (Huebner, 2013; Hutchins, 1995; Goldstone; Gureckis, 2009; Goldstone; Theiner, 2017; Theiner; O’Connor, 2010; Barnier et al., 2008; List; Pettit, 2011; Schwitzgebel, 2015). Essa não parece ser uma coincidência, o reconhecimento da natureza descentralizada do processamento cognitivo humano dá um ar de maior plausibilidade para a tese de que tais entidades distribuídas

estados psicológicos a grupos (Knobe; Prinz, 2008; Systma; Machery, 2009, 2010; Arico, 2010, Hubner et al., 2010). Note, também, que uma imagem naturalista completa deve levar em consideração resultados das Ciências Sociais, pois a questão que temos em mãos se divide em dois componentes centrais: “O que é cognição?” e “Grupos têm estados cognitivos ou envolvem-se em processamento cognitivo?” Aqui eu me concentrarei na primeira questão e, portanto, minha ênfase será na Filosofia das Ciências Cognitivas e não na Filosofia das Ciências Sociais.

como tribunais, corporações e sociedades possam ter estados cognitivos genuínos. De fato, alguns autores discutem as duas questões explicitamente em paralelo (Hutchins, 1995; Wilson, 2004; Tollefsen, 2006; Theiner; Allen; Goldstone, 2010).

A questão que naturalmente surge diante disso é se essas duas linhas de pesquisa - sobre a cognição distribuída nos humanos individuais, de um lado, e sobre processos de grupos, de outro lado - investigam os mesmos tipos naturais, estados, propriedades ou processos⁵, como por exemplo *crenças*, *memória* e *cognição*. Se grupos parecem ter memórias próprias, seriam esses estados do mesmo tipo natural que os estados de humanos individuais aos quais nos referimos como “memórias”? O que significa haver tanto um tipo abrangente de *memória* como também uma espécie particular dela? Se há cognição de nível do grupo, será ela o mesmo *tipo* de cognição com a qual humanos individuais se envolvem? E, alguém poderia se perguntar, o que significaria haver espécies diferentes no gênero *cognição*? Quão profundas são as relações de igualdade de tipo entre os casos individuais e de grupo e como alguém deve proceder ao tentar determinar essas igualdades?

Aqui está uma prévia do que vem a seguir. Na Seção 2, eu examino dois argumentos a favor do realismo sobre estados de grupos. Eu considero ambas linhas de raciocínio inconclusivas, em grande medida porque elas se desenvolvem sem uma abordagem sobre igualdade de tipo natural e, portanto, deixam a questão central não respondida. Na Seção 3, eu articulo uma abordagem do tipo de evidência que poderia sustentar afirmações de igualdade (ou diferença) de tipos naturais (incluindo, de modo mais significativo, tipos finos (*thin*), genéricos ou superordenados)

⁵ Para os presentes propósitos, usarei esses termos intercambiavelmente, todos eles fornecem uma maneira de se referir às juntas da natureza que as ciências tentam identificar.

entre casos. Em outro lugar eu chamei isso de abordagem de “ajuste e extensão” (“*tweak-and-extend*”) (Rupert, 2013). Para aplicar o teste do ajuste e extensão, devemos ter em mãos modelos dos dados que devem ser explicados pelas instâncias hipotetizadas do tipo em questão. Portanto, na Seção 4, eu apresento o que eu acredito que seja o melhor modelo do tipo genérico de *cognição*, ao menos da forma como ela aparece em casos humanos. Se essa teoria sobre cognição estiver correta, ela remove uma barreira potencial para cognição genuína de nível de grupo; pois, essa teoria permite que um indivíduo com estados cognitivos seja descentralizado. A seção 5 dá continuidade a essa ideia, a partir do material apresentado nas Seções 3 e 4 e desenvolve uma versão mais definitiva da proposta de que grupos têm estados cognitivos que são de um tipo diferente de cognição daquela dos humanos individuais. A Seção 6 considera objeções e conclui com uma nota pessimista.

2. Realismo, redução e similaridade funcional

Uma teoria é realista com respeito a cognição de grupo se e somente se, de acordo com essa teoria, (a) em um número significativo de casos, um grupo de humanos, enquanto uma entidade singular, têm estados cognitivos próprios e (b) tais estados são do mesmo tipo relevante, *qua* estados cognitivos, que os estados de humanos individuais. O tipo de realismo que tenho em mente não repousa em equivocação, metáfora, atribuição “como se” ou o uso de homógrafos ou termos de tipo “meramente nominais”. O aspecto radical dessa teoria surge a partir de seu realismo audacioso sobre muitos grupos concretos: quando a sentença “A Inglaterra acredita que sair da União Europeia será vantajoso” é verdadeiramente asserida, de acordo com o realista essa sentença se torna verdadeira por um estado de uma entidade

singular, a saber, o Reino Unido ter uma crença particular do mesmo tipo natural que as crenças mantidas por humanos individuais (seja a favor ou contra o *Brexit* ou a respeito de qualquer outro assunto).

Na presente seção, examino dois argumentos a favor do realismo. O primeiro tem um sabor de argumento anti-reducionista (no entanto, veja a Seção 6 para ressalvas sobre essa caracterização). O segundo foca na similaridade (*similarity*) do processamento cognitivo em grupos com o processamento cognitivo em humanos individuais. Nenhum argumento é claramente bem sucedido, ou assim argumentarei. Discussões sobre os pontos fracos dos argumentos põe em evidência o potencial de julgamento de uma abordagem sobre a relação de *ser do mesmo tipo natural que*.

2.1 Argumentos baseados em princípios naturalistas de compromisso ontológico

O primeiro argumento aparece em um artigo recente de Christian List (2018) que se concentra predominantemente na possibilidade de uma consciência fenomênica de grupo, um tópico que deixarei de lado aqui. Como uma maneira de enquadrar a discussão, List resume aquilo que toma como sendo o argumento mais forte a favor do realismo sobre cognição de grupo⁶, um argumento que ele ajudou a desenvolver em uma série de publicações anteriores, talvez mais notavelmente em um livro escrito em coautoria com Philip Pettit (List; Pettit, 2011):

Premissa 1: Nossas melhores teorias científico-sociais de certos fenômenos sociais - por exemplo, nossas melhores teorias sobre o comportamento das firmas no mercado - atribuem agência baseada em crenças e desejos do tipo funcionalista para (alguns dos) os coletivos envolvidos, frequentemente os representando como agentes no sentido das teorias dos jogos e da decisão.

⁶ List é um realista sobre cognição de grupo, mas não sobre consciência fenomênica de grupo.

Conclusão intermediária: De acordo com uma definição naturalista do compromisso ontológico, essas teorias são, portanto, comprometidas ontologicamente com grupos agentes.

Premissa 2: Nós devemos, ao menos falivelmente, aceitar os compromissos ontológicos de nossas melhores teorias científicas em qualquer domínio.

Conclusão: Nós devemos, ao menos falivelmente, aceitar o compromisso das nossas melhores teorias científico-sociais sobre grupos agentes. (List, 2018, p. 298).

Para tornar mais claro, aceitar esses compromissos à primeira vista, compromete alguém, ao menos provisoriamente, com a existência de grupos agentes genuínos os quais têm suas próprias crenças e desejos (ou julgamentos e preferências, como é frequentemente colocado em List e Pettit [2011]).⁷

Dada a referência de List a teorias científico-sociais e seu apelo ao princípio naturalista do compromisso ontológico, o leitor pode assumir com segurança que ele tem em mente tipos naturais ou cientificamente respeitáveis. Se a *Premissa 1* for verdadeira e o sucesso da teoria é genuíno, então os tipos em questão são, de fato, cientificamente respeitáveis. (Em alguns casos, nossos melhores modelos não são muito bem-sucedidos; mas no presente contexto, a caridade exige que entendamos que a melhor teoria implica que ela é, pelo menos, moderadamente bem-sucedida - o bastante para autorizar o compromisso ontológico).

⁷ Isso representa apenas um dos dois argumentos realistas desenvolvidos por List e Pettit - o argumento da ciência social positiva (List; Pettit, 2011, p. 10). List e Pettit também constroem um caso normativo a favor do realismo que se baseia em intuições como a de que toda ação boa ou má requer um rateio apropriado de uma quantidade correspondente de elogio ou culpa; da forma como List e Pettit veem as coisas, em alguns casos, a culpa ou elogio apropriadamente distribuído entre os membros individuais de um grupo não totalizam uma quantidade suficiente dada a natureza do ato em questão e, portanto, um grupo agente existe como uma coisa que deve ser responsabilizada e recebe o restante de elogio ou da culpa (p. 166). Tais intuições normativas parecem claramente fora de lugar no presente contexto, cujo tema é, grosso modo, "o que as Ciências Cognitivas nos dizem sobre o realismo", pois tais intuições não têm força probatória nas Ciências Cognitivas, em vez disso, elas existem apenas como fenômenos, isto é, como dados; a produção dos quais deve ser modelada usando as ferramentas padrão das Ciências Cognitivas; e não há razão para pensar que os melhores modelos disponíveis nas Ciências Cognitivas para a produção de tais intuições incluirão fatos normativos (no sentido filosófico) como elementos.

Uma preocupação central no presente contexto é a questão de se crenças e desejos como caracterizados pelas teorias científico-sociais em questão são os mesmos tipos-de-estados cientificamente respeitáveis que os tipos-de-estados em que humanos tipicamente estão quando eles têm aquilo que é comumente chamado de “crenças” ou “desejos”. Não está garantido que todas as aplicações de “crença” e “desejo” refiram ao mesmo tipo natural, assim como as aplicações de “jade” não estão (Putnam, 1975). Considere que grupos típicos e humanos individuais típicos diferem enormemente em aspectos que são, ao menos à primeira vista, psicologicamente relevantes. Tome, por exemplo, questões que dizem respeito à vida doméstica e interações interpessoais. Grupos não postam perfis de namoro em sites da Web e não combinam encontros para tomar um café ou um drinque. Grupos não colocam seus filhos para dormir, assegurando-se de trocar as fraldas de seus grupos-bebês antes disso. Humanos individuais têm padrões característicos de desenvolvimento e experiências que grupos não têm, por exemplo, a experiência de aprender uma música sobre o alfabeto de sua primeira língua ou de passar a entender aspectos de dinâmicas de grupo ao participar de atividades esportivas organizadas. E assim por diante, ao longo de uma grande variedade de assuntos humanos. Não seria surpreendente, portanto, descobrir que os estados cognitivos operantes na produção de uma vida distintivamente humana sejam de um tipo diferente dos estados do nível de grupo que produzem, por exemplo, decisões de tribunais, mesmo que esses estados sejam algumas vezes chamados de “crenças” e “desejos” e mesmo que eles desempenham um papel explanatório-causal importante nas Ciências Sociais, não sendo de forma alguma desinteressantes.⁸

⁸ O ponto aqui é que estados de grupo podem ser de enorme interesse científico e filosófico mesmo se grupos não tenham estados cognitivos ou psicológicos - ou, sendo mais cauteloso, não tenham estados

Como essa questão deve ser resolvida? Uma abordagem convincente a respeito das igualdades e diferenças dos tipos naturais - ou do tipo de evidência que alguém deve ter para a igualdade ou diferença de dois membros de um tipo natural - forneceria um caminho promissor. Esse será o tema da Seção 3, abaixo.

Para ser justo, no entanto, List pode considerar que a questão sobre a igualdade do tipo tenha sido respondida. Afinal de contas, List e Petit se esforçam para afirmar e explicar uma teoria sobre o que é ser um agente: uma entidade é um agente se e somente se ela tem estados representacionais, estados motivacionais e a habilidade de combiná-los de maneiras apropriadas para agir sobre seu ambiente (List; Pettit, 2011, p. 20). Essencialmente, isso quer dizer que algo é um agente se e somente se ele tem uma psicologia de crenças e desejos, comportando-se de uma forma que obedece, e pode ser prevista e explicada pela, a aplicação do silogismo prático.

No entanto, isso é muito esquemático para ser satisfatório. Essa proposta explora a familiaridade que o leitor tem com a psicologia popular, isto é, com nossas valiosas práticas explicativas cotidianas na nossa interação com coespecíficos; mas ela não enfrenta as questões difíceis sobre a igualdade genuína de tipos naturais. Particularmente, ela não estabelece que (a) os tipos da psicologia popular são tipos naturais (isto é, se estados representacionais e motivacionais, *da forma como são concebidos popularmente*, aparecem em qualquer teoria científica razoavelmente bem-sucedida para autorização de compromisso

do mesmo tipo natural que os estados humanos identificados quando utilizamos termos cognitivos, psicológicos ou mentalistas. Eu incluo essa qualificação porque minha preocupação imediata não é a de identificar estados cognitivos, mentais ou psicológicos *reais*, onde ser o item genuíno for pensado como que carregando algum tipo de privilégio ou estatuto normativo, somente na afirmação do realista é que os estados de interesse são do mesmo tipo entre os casos humanos e de grupo.

ontológico)⁹ ou (b) que os estados mencionados na Premissa 1 de List são do mesmo tipo que a espécie de estados populares confirmados pelas Ciências Cognitivas (ou são do mesmo tipo que os estados que as Ciências Cognitivas atribuem no lugar dos tipos-de-estados da psicologia popular). As Ciências Cognitivas produziram teorias sobre estados motivacionais e representacionais, mas, geralmente, elas caracterizam estados, processos e mecanismos que provavelmente não aparecem em grupos e, portanto, o sucesso dessas teorias dificilmente mostra que uma propriedade fina (*thin*) como *ser um estado motivacional* desempenha uma papel explanatório-causal robusto.

Permita-me abordar essa questão por outro ângulo. List e Pettit configuram suas discussões sobre tipos-de-estados em um contexto teórico, um contexto indicado pela Premissa 1 de List. Eles são funcionalistas sobre estados mentais (Putnam, 1967; Lewis, 1980; Shoemaker, 1981). List e Pettit não podem ser criticados por isso; a abordagem funcionalista para tipos-de-estados mentais é a posição predominante na Filosofia da Mente. Entretanto, essa perspectiva funcionalista fornece também uma estrutura perspicua dentro da qual podemos identificar a lacuna no argumento de List.

Na teoria funcionalista, ser um dado tipo de estado mental é ser um estado que desempenha um papel causal-funcional distintivo na economia total de estados de um sujeito fazendo a mediação entre *inputs* e *outputs*. Para pertencer a espécie de tipos-de-estados que um humano instancia é necessário ser um estado que ocupa o mesmo lugar que aquele ocupado pelo estado humano correspondente em um construto interno massivamente interconectado, envolvendo esperança, percepção, medo,

⁹ Reconheço (ainda que permaneça cético sobre) as razões de Fodor para pensar que a psicologia computacional reivindica uma versão ontologicamente robusta da teorização da psicologia popular (Fodor, 1987, cap. 1; cf. Churchland, 1981).

memória, sonho, inferência, etc. (ou seus descendentes nas Ciências Cognitivas). Quão detalhada deve ser essa rede é uma questão controversa (Block, 1978), mas esse tanto é claro: List deve identificar uma concepção funcionalista de crença e desejo suficientemente refinada para que os estados em questão desempenhem um papel explanatório-causal significativo (ou apareçam como elementos não estranhos em nossos melhores, razoavelmente bem-sucedidos, modelos de comportamento humano), mas não refinada demais, de modo que eles claramente não seriam possuídos pelos grupos típicos de interesse.¹⁰

Assuma a posição razoável de que as sentenças Ramsey (Lewis, 1970) para a psicologia humana, da forma como as Ciências Cognitivas (e não a psicologia popular) apresentam, contenham certa quantidade modesta de detalhes. Então, qual seria a relação entre elas e modelos científico-sociais? É extremamente duvidoso que teorias científico-sociais Ramsificadas - uma teoria sobre organizações políticas ou uma teoria sobre organizações jurídicas - tenham algo remotamente parecido com a estrutura de uma sentença Ramsey para a psicologia humana. Nesse caso, dada a natureza holística da individuação de estados funcionalistas, humanos e os grupos em questão não compartilham nenhum dos estados relevantes. No mínimo, deveríamos requerer uma abordagem positiva sobre a igualdade dos tipos naturais - para além de uma comparação rasa de nossas explicações do comportamento de grupos com as explicações da psicologia popular em termos de crenças e desejos - para sustentar a afirmação realista de que os tipos naturais em questão são os mesmos.¹¹

¹⁰ Essa linha argumentativa é paralela ao dilema apresentado em Rupert (2004, seções V-VIII) no contexto do debate sobre a extensão da mente e da cognição no nível individual (Clark; Chalmers, 1998).

¹¹ O processo de Ramsificação fornece uma ferramenta formal de esquematização das teorias psicológicas. Uma sentença Ramsey faz isso através da amalgamação ou costura de muitos modelos fragmentários bem-sucedidos (isto é, modelos de conjuntos de dados específicos ou modelos de

O que tende a tornar o problema em questão obscuro é a natureza fragmentária de muitos modelos influentes das Ciências Sociais que interessam a List, incluindo, por exemplo, as estruturas explanatórias típicas das teorias dos jogos e da decisão, aquelas discutidas em conexão com o Dilema dos Prisioneiros (List; Pettit, 2011, p. 107-108) ou os mecanismos para manutenção da consistência geral em decisões do nível do grupo quando confrontadas com o Dilema Discursivo (p. 45-46). É claro que muitas das estruturas explanatórias empregadas na modelagem cognitiva também são meros fragmentos de uma teoria psicológica, mas o Santo Graal das Ciências Cognitivas é a construção de uma arquitetura (e.g., SOAR, LEABRA, ACT-R ou GPS), dentro de um contexto no qual os vários fragmentos se unem de forma coerente como partes de um único agente (idealmente autônomo) - onde o analisador semântico de linguagem está funcionalmente entrelaçado com os mecanismos de reconhecimento de face, e assim por diante. Considere a afirmação de que uma empresa pode agir racionalmente no mercado, uma afirmação que se baseia em uma certa concepção do que é optimal por meio de uma função que relaciona *inputs* e *outputs*. Tomada fora do contexto de uma Ramsificação da psicologia humana como um todo, isso pode parecer refletir a operação dos mesmos tipos de estados que estão em operação quando um ser humano individual otimiza sua performance com respeito à, digamos, suas finanças domésticas. Entretanto, quando configurada dentro do contexto da Ramsificação completa da psicologia individual e a Ramsificação completa de uma teoria da estrutura corporativa, parece que não haverá mérito na

famílias de conjuntos de dados relacionados). Tipos-de-estados abrangentes - tais como crença - podem ser explicitados nos modelos relevantes e, portanto, na sentença Ramsey construída a partir deles; ou tais tipos podem ser apenas implicitamente definidos, digamos, por padrões estruturais comuns detectáveis em várias partes da sentença Ramsey em questão. De qualquer modo, a questão permanece aberta se os tipos, assim definidos, são compartilhados por grupos e humanos individuais.

afirmação da igualdade de tipo, dado que a sobreposição entre estruturas provavelmente será muito pequena e fragmentária. A explicação particular do nível de grupo em questão depende de uma coleção de estados em interação; cujo comportamento pode corresponder a uma pequena coleção de estados da Ramsificação da psicologia humana e, mesmo assim, somente dentro de um faixa extremamente limitada. A correspondência de um pequeno fragmento ao longo de uma faixa extremamente limitada de casos dificilmente será uma base para afirmar a igualdade de tipo natural - ao menos não o será se alguém assumir, como List o faz, uma concepção funcionalista dos estados mentais. Qualquer um que tente avaliar o argumento de List deveria querer, no mínimo, uma abordagem mais explícita do tipo de evidência que sustentaria a sua afirmação de igualdade dos tipos naturais entre indivíduos e grupos (isto é, do tipo de evidência que sustentaria a afirmação de que os vários usos de “crença”, etc. em questão referem ao mesmo tipo natural nos casos de humanos e de grupos) e uma aplicação desse critério à questão em pauta.¹²

¹² List e Pettit não se concentram exclusivamente em estados individuais de agentes, tais como crenças e desejos. Eles frequentemente enfatizam propriedades da psicologia do agente como um todo, tais como a consistência e coerência ao longo do tempo. Eles argumentam que para construirmos instituições de nível do grupo que sirvam aos nossos interesses, nós devemos estruturar cuidadosamente essas instituições - por meio da codificação do uso de certos métodos de agregação de julgamentos e preferências de indivíduos humanos - de uma forma a dotar grupos com tais propriedades como consistência, coerência, completude (onde ser completo é ser tal que, para qualquer proposição relevante, o grupo a endossa ou endossa a sua negação; List; Pettit, 2011, p. 53), a ideia, em parte, é que tais propriedades devem ser satisfeitas por grupos para que uma psicologia de crença e desejo seja proficuamente aplicável a eles.

No entanto, List e Pettit fazem muito pouco para sustentar a afirmação de que as propriedades em questão são, de fato, possuídas por humanos individuais. Eles apelam para concepções intuitivas dos papéis que tais propriedades desempenham na psicologia humana e para afirmações sobre como nós gostaríamos que nossas instituições funcionassem, sem argumentar que as propriedades globais que poderemos criar nos grupos são do mesmo tipo natural que as propriedades globais dos sistemas cognitivos de humanos individuais. Ainda que uma concepção vaga de racionalidade provavelmente possa ser aplicada para humanos e para certos grupos, a questão permanece aberta sobre o quão fino é o tipo compartilhado em questão e se um tipo tão fino desempenha qualquer trabalho explanatório-causal.

2.2 Argumentos baseados na similaridade

Muitos autores argumentaram a favor do realismo de mente de grupos fazendo comparações diretas com casos de humanos individuais. Foi afirmado que o processamento de nível de grupo envolve a complexa transformação computacional de representações (Hutchins, 1995) que amplamente é tida como a base da cognição humana ou que esse processamento inclui a operação de algoritmos ou mecanismos - tais como inibição lateral (Goldstone; Theiner, 2017) - que desempenha um papel significativo em modelos de processamento humano. Pareceu a muitos que de tais similaridades se segue que grupos têm o mesmo tipo de estados mentais ou cognitivos que humanos. (Huebner, 2008, p. 109, 2013; Rovane, 2014, p. 1663; Tollefsen, 2006, p. 144; Theiner; O'Connor, 2010, p. 89; Theiner et al., 2010, p. 382-383).

Tais argumentos parecem mais convincentes quando eles apelam para similaridades finas entre humanos e grupos, similaridades na variedade de propriedades e processos de interesse no estudo científico da cognição humana. Se podemos identificar similaridades com respeito aos próprios processos, estruturas e tipos que fazem o trabalho explanatório-causal nos modelos de dados das Ciências Cognitivas envolvendo performance humana inteligente, podemos estar confiantes de que encontramos tipos naturais comuns da espécie relevante.

No entanto, essa é uma tarefa árdua, improvável de ser cumprida. Como ilustração, considere a proposta bastante influente de que a arquitetura da mente humana contém módulos especializados. Na proposta de Fodor (1983), a arquitetura humana consiste em (pelo menos) (a) transdutores periféricos, (b) módulos de domínio específico (e.g., sistemas sensoriais), (c) uma unidade de processamento central que raciocina holisticamente e (d) sistemas de *output*. Esse é um esquema

inicial. Mas, certamente, cientistas cognitivos passam a maior parte de seu tempo preenchendo detalhes desse esquema arquitetônico. Por exemplo, com relação ao sistema de *output*, foi afirmado que o sistema motor humano inclui circuitos emuladores; uma parte de tal circuito recebe uma cópia de um comando motor exterior (uma, assim chamada, cópia eferente) e executa uma simulação muito rápida que pode, por exemplo, prever se o membro sendo movido pousará onde deve pousar, em tempo suficiente para que um comando motor compensatório seja enviado, caso o membro pareça não estar na trajetória apropriada (Grush, 1997). No que diz respeito ao esmiuçamento dos detalhes do processamento cognitivo humano, essa é apenas a ponta do iceberg. Na medida em que os detalhes aumentam, a ideia de que o processamento de grupo possa ter a mesma estrutura do processamento humano - ao mesmo nível de refinamento que geralmente interessa ao cientista cognitivo - parece inacreditável. Perfis detalhados do funcionamento cognitivo humano ocupam milhões de páginas, tomos após tomos, artigo após artigo, catalogando os frequentemente excêntricos e altamente específicos trabalhos sobre, por exemplo, a memória humana e a visão humana. Se alguém se concentrar nessas espécies de estados, processos e perfis causais que interessam ao cientista cognitivo, a busca pela similaridade parece um beco sem saída.

Ao mesmo tempo, na medida em que esses detalhes se acumulam, começamos a nos incomodar com o chauvinismo de tudo isso; talvez a arquitetura humana - com todas as suas formas de processamento excêntricas e desajeitadas (*kludgy*) - seja suficiente para a cognição. Mas, será ela necessária? Um grupo precisaria trabalhar *exatamente dessa forma* para ter cognição? Porque não dizer, ao invés disso, que alguns grupos têm cognição, mas que ela é de uma espécie diferente da

cognição humana?¹³ Para tornar essa ideia convincente, no entanto, será requerida uma abordagem sobre quando dois tipos são espécies de um mesmo tipo genérico e abrangente de *cognição*.

3. A abordagem do ajuste e extensão de compartilhamento de tipo

O que, então, distingue um tipo genérico, mas natural, de um mero tipo nominal? Minimamente, várias espécies de um tipo natural genérico mantêm uma semelhança (*resemblance*) de família entre si (Wheeler, 2011), mas não é qualquer semelhança de família que será suficiente; deve existir, eu argumento, uma unidade explanatório-causal para as várias instância de um tipo. Afirmações de igualdade de tipo são mais convincentemente legitimadas por certos padrões em modelos dos fenômenos de interesse: nossos melhores modelos das formas pelas quais as várias espécies produzem instâncias do *explananda* relevante devem ter um número significativo de sobreposição de elementos e relações entre esses elementos. Essa parece ser a ordem do dia na maioria das ciências; um modelo inicial (geralmente simples) de algum fenômeno paradigmático é (suficientemente) bem-sucedido e, então, fenômenos relacionados são modelados por meio do “ajuste” do modelo inicial - termos e elementos do modelo são adicionados, valores de parâmetros são ajustados, etc. - e a aplicação da família de modelos em questão é subsequentemente estendida e, talvez, aprimorada (isto é, tornam-se mais precisos enquanto modelos para sistemas aos quais seus primos anteriores ao ajuste foram aplicados). Se um fenômeno que nós pensávamos ser

¹³ A esse respeito, a presente discussão se encaixa em um longo debate sobre funcionalismo e chauvinismo no nível individual (Block, 1978). Sobre o papel que essa questão desempenhou em debates sobre cognição distribuída e mente estendida, veja Rupert (2004, 2013), Clark (2008), Sprevak (2009) e Wheeler (2011).

parecido com outros acaba não sendo sensível a esse tratamento de ajuste e extensão, ele é, e deveria ser, tratado como um tipo de fenômeno diferente; nós devemos concluir que, enquanto se aguardam novos desenvolvimentos, não há um tipo genérico que inclui todos os casos relevantes.

Como poderia a abordagem do ajuste e extensão funcionar na prática das Ciências Cognitivas? Considere, por exemplo, a apresentação esquemática da ideia de um modelo conexionista em Rumelhart, Hinton e McClelland (1986); ele contém, dentre outras coisas, unidades com níveis individuais de ativação, dispostas em camadas e que podem transmitir a ativação adiante para unidades nas camadas adjacentes, cujos níveis de ativação são determinados por uma função que considera a ativação (ou inibição) recebida como um *input*, e assim por diante. Os autores apresentam esses elementos básicos e, com respeito a cada um deles, descrevem efeitos diferenciais de, por exemplo, diferentes configurações de parâmetros possíveis ou escolhas de funções de ativação.

Diferentes formas de preencher o esquema produzem diferentes tipos de comportamentos nas redes resultantes, introduzindo a possibilidade de que alguém construa uma próspera família de modelos inter-relacionados que explique uma variedade cada vez maior de formas de comportamento inteligente (p. 46).

Então, ser um modelo conexionista é ser um modelo que instancia a receita de Rumelhart et al. ou ser um modelo que mantém uma relação de ajuste e extensão com instanciações do esquema de Rumelhart et al. E nossa melhor evidência de que dois tipos de processos são ambos instâncias de *processamento conexionista* - isto é, que eles compartilham o tipo natural em questão - é que nossos melhores modelos dos dois processos são ambos modelos conexionistas. Essa abordagem elimina

certas possibilidades: alguém não pode, sem um contorcionismo excessivo, começar com um modelo proto-PDP (uma instanciação direta do esquema de Rumelhart et al.) e transformá-lo por meio de ajustes em uma tabela de consulta (*look-up table*).

Eu proponho a abordagem de ajuste e extensão como uma explicação da evidência relevante para o julgamento sobre a igualdade de tipo, uma abordagem que é condizente com o trabalho realizado em todas as ciências. Considere, por exemplo, *ser um oscilador harmônico*. Começamos com uma equação simples que descreve o comportamento de um pêndulo idealizado; na tentativa de modelar uma maior variedade de sistemas, adicionamos um termo para fricção e, então, uma função de força, e assim por diante. O fato de esta família de modelos “ajustados e estendidos” ter aplicações frutíferas para uma variedade de sistemas físicos é forte evidência de que esses sistemas instanciam um tipo comum - são *osciladores harmônicos*.¹⁴

Uma apresentação e defesa completa da abordagem de ajuste e extensão excederia as limitações de espaço. Aqui me contentarei com a discussão de uma possível objeção sobre os limites entre modelos que são relacionados por ajuste e extensão, de um lado, e aqueles que não são assim relacionados, de outro. Alguém poderia se perguntar como essa proposta lida com uma série muito longa de ajustes. Seria relativamente fácil construir, ao estilo de um argumento incremental do tipo sorites, uma família de modelos inter-relacionados por ajustes onde dois membros, analisados em comparação, pareceriam significativamente não relacionados um com o outro?¹⁵ Talvez o passo razoável mais simples nesses casos seja o de deferir para a intuição dos

¹⁴ Para outro exemplo, veja a discussão de Chemero do modelo Haken-Kelso-Bunz e suas várias extensões e refinamentos (2009, p. 85–96).

¹⁵ Agradeço Jackson Kernion por me pressionar neste ponto.

especialistas e profissionais, para que nos digam quando uma série de ajustes produz um resultado “irreconhecível”.¹⁶

Idealmente, no entanto, deveríamos buscar uma forma mais substantiva de separar uma similaridade aparente de uma similaridade que genuinamente indica igualdade de tipo. Isso pode envolver a identificação de detalhes das variações dos dados que são explicados pelos componentes compartilhados dos dois modelos (isto é, qualquer coisa que permaneça comum a eles depois do ajuste requerido para construir um modelo a partir do outro). O método pode assumir a seguinte forma: selecione dois modelos candidatos; identifique seus elementos ou estruturas compartilhadas; diversifique os sistemas aos quais esses modelos estavam orientados de forma bem-sucedida no passado ao longo da dimensão em que os componentes compartilhados dos modelos foram mapeados nos casos de aplicações anteriores bem-sucedidas; veja se variações correspondentes nos componentes compartilhados dos modelos explicam a variação em outras quantidades mensuráveis do sistema alvo (variações que acompanham alterações em parte do sistema alvo que corresponde aos componentes compartilhados dos modelos). A ideia é descobrir evidência de que aquilo que é compartilhado pelos dois modelos candidatos desempenha um papel explanatório-causal ou se isso é, de alguma forma, incidental. Se esses elementos desempenham um papel explanatório-causal, então nós colocamos os modelos candidatos em uma mesma família, mesmo se, *prima facie*, a proporção dos processos de ajuste e extensão acabe tornando os modelos primos distantes.

¹⁶ Embora seja possível se perguntar até que ponto tais julgamentos vão variar com o contexto específico, com, por exemplo, o propósito que o especialista tem em mente quando lhe é perguntada a questão relevante (cf. Block, 1997).

Note que mesmo quando um elemento compartilhado desempenhar trabalho explanatório-causal, permanece a questão de como melhor individuar o tipo natural para o qual evidência foi obtida. Os aspectos relevantes do sistema alvo (os aspectos que correspondem aos componentes compartilhados em questão, bem como as alterações que foram explicadas pelas mudanças nos valores dos componentes compartilhados) devem guiar nossa caracterização do tipo. Por exemplo, se nós tratarmos os componentes compartilhados dos modelos como indicativos da presença de um tipo genérico de *cognição*, então os dados em questão devem ser relevantes para aquilo que nós consideramos como fenômeno cognitivo, e não, digamos, para a massa ou tamanho do sistema alvo.

O que apresentei é, por necessidade, um esboço incompleto. Tenha em mente, no entanto, que meu propósito não é apresentar uma teoria do que é ser um tipo genérico, mas articular um quadro mais estruturado dentro do qual se pode tentar construir um argumento naturalista a favor do realismo.

4. O que é cognição?

Se grupos têm estados cognitivos - isto é, possuem estados do mesmo tipo natural que os estados cognitivos humanos - depende parcialmente de fatos empíricos sobre a própria cognição. Cognição é um tipo científico, hipotetizado pelos cientistas relevantes para explicar (o que nós consideramos ser) um domínio particular de fenômenos que dizem respeito, em um primeiro momento, ao comportamento de humanos individuais: conversas em tempo real, padrões de similaridade no tratamento de objetos (reidentificação), a produção de obras de arte, a formulação e testagem de teorias

científicas, jogar xadrez, desempenho em exames de compreensão de leitura de texto, etc.¹⁷

O que é, então, cognição? Em outros trabalhos (Rupert, 2004, 2009, 2010), eu argumentei que virtualmente todas as formas bem-sucedidas de modelagem cognitiva - computacional, biologicamente pura, baseada em robótica, conexionista e dinamicista - distinguem, de um lado, arquiteturas relativamente integradas e persistentes e, de outro lado, contribuintes causais transientes que, junto com características da arquitetura, produzem comportamento inteligente (cf. distinção de Wilson [2002] entre sistemas obrigatórios e facultativos) e que, além disso, tais elementos nos fornecem as melhores pistas para a distinção entre causas genuinamente cognitivas para a produção de comportamento inteligente e outras causas contribuintes. Pense nisso como uma inferência a melhor explicação (disponível) realizada duas vezes. Primeiro, que alguns aspectos centrais e distintivos do processamento cognitivo são normalmente instanciados dentro dos limites de organismos individuais explica porque vários modelos que focam em indivíduos têm sido bem-sucedidos. Segundo, que a arquitetura persistente ser a coisa distintivamente cognitiva melhor explica por que ela perpassa as várias formas de modelagem de nível individual bem-sucedidas.

Podemos dizer algo mais preciso sobre a natureza integrada do sistema que aparece dentro da pele, qualquer coisa que ilumine seu papel como um sistema cognitivo, isto é, um sistema que flexivelmente produz uma ampla gama de formas de comportamento inteligente? Em

¹⁷ Essa abordagem pode parecer uma petição de princípio no presente contexto, mas ela representa uma metodologia padrão nas ciências. Se o processo de investigação funcionar apropriadamente, então, no curso normal da pesquisa científica, quais dos vários fenômenos se revelam do mesmo tipo - isto é, quais são, de fato, produzidos por processos do mesmo tipo - emergirão, independentemente das inclinações iniciais dos pesquisadores para agrupar os fenômenos de uma forma ou de outra.

trabalhos anteriores sobre o tema (Rupert, 2009, 2010, 2011, 2013), eu propus que um sistema cognitivo consiste em um coleção de mecanismos que co-tribuem em subconjuntos sobrepostos para a produção de uma ampla gama de formas de comportamento inteligente, como também propus um critério matemático que explicaria a condição “em subconjuntos sobrepostos” (Rupert, 2009, 2010). Esse critério matemático é, em essência, uma teoria da integração. Apesar de ter sido motivada por uma tentativa de caracterizar o sistema cognitivo interno, o critério é neutro quanto à localização; ela diferencia dois tipos de contribuintes causais, onde quer que eles apareçam, para a produção de comportamento inteligente.

Aqui está o que eu agora chamarei de teoria da “probabilidade condicional de co-contribuição” para a integração cognitiva, ou PCC, refinada de modo a clarificar sua estrutura. Tenha em mente que, embora a descrição a seguir soe procedural - como se ela fosse uma receita para levar a cabo uma construção - ela objetiva caracterizar a propriedade de integração cognitiva:

1. Para qualquer sujeito, em qualquer tempo, modele cada subconjunto não unitário dos mecanismos que possuem contribuição causal distintiva para a produção de qualquer forma de comportamento inteligente.
2. Para cada subconjunto desses, relativo a cada forma de comportamento inteligente, há, para cada um de seus subconjuntos, uma probabilidade de que ele seja um contribuinte causal para a produção daquela forma de comportamento condicional a cada um dos membros do complemento do conjunto de contribuintes causais.

3. Faça um ordenamento de todas essas probabilidades condicionais.
4. Faça o recorte natural entre probabilidades altas e baixas. (Se ser um sistema integrado é um tipo natural, e a proposta sendo oferecida está no caminho certo, deveríamos esperar que apareça uma lacuna estatisticamente significativa como essa.)
5. Para cada mecanismo que aparece na lista de conjuntos com probabilidades condicionais altas (isto é, os conjuntos acima da lacuna referido no Passo 4), conte o número de vezes que esse mecanismo aparece e ordene os mecanismos de acordo com isso (isto é, de acordo com o número de vezes que eles aparecem acima da lacuna da lista produzida pelo Passo 4).
6. Uma lacuna estatisticamente significativa separará os mecanismos que aparecem frequentemente nessa segunda lista daqueles que não aparecem frequentemente.
7. O sistema cognitivo integrado abrange todos e somente aqueles mecanismos que aparecem acima da lacuna da segunda lista.

Apresentada dessa maneira puramente formal, as implicações da PCC podem permanecer obscuras. Portanto, considere um exemplo. O sujeito típico é frequentemente habilidoso em desviar de obstáculos na medida em que se move e um mecanismo de detecção de bordas quase certamente contribuiu causalmente para tal comportamento. Um mecanismo que computa a distância a partir de disparidades da retina também terá contribuído para o desvio dos obstáculos pelo sujeito típico, assim como um mecanismo que calcula o formato a partir da detecção de sombras (Marr, 1982). Com respeito ao desvio de obstáculos, muitos mecanismos adicionais contribuíram, por exemplo, vários mecanismos de controle motor. Para manter as coisas relativamente simples, vamos

adicionar apenas um desses mecanismos de controle motor para o conjunto de mecanismos sendo considerados. O conjunto resultante de quatro mecanismos permite a possibilidade de seis conjuntos de dois membros, quatro conjuntos de três membros e um conjunto de quatro membros. Para cada conjunto de dois membros, duas probabilidades condicionais são relevantes: a contribuição condicional do primeiro mecanismo sobre o segundo e vice-versa; isso resulta em um total de doze itens na lista ordenada construída no Passo 3 da PCC. Para cada conjunto de três membros, há seis probabilidades condicionais relevantes: a contribuição condicional de cada mecanismo individual sobre os outros dois e cada combinação da contribuição de dois mecanismos sobre o terceiro; isso resulta em um total de vinte e quatro itens adicionais na lista ordenada construída no Passo 3 da PCC. Para conjuntos de quatro membros, há quatorze probabilidades condicionais relevantes (as quais representarão mais quatorze itens na lista ordenada em questão). Para qualquer um dos quatro, devemos incluir a probabilidade de sua contribuição condicional sobre a contribuição dos três restantes, e vice-versa, o que resulta em oito itens. Os subconjuntos restantes dos conjuntos de quatro membros são pares, assim como os complementos em todos esses casos. Então, escolha qualquer um dos seis pares possíveis e há uma probabilidade condicional de sua contribuição dado que seu complemento esteja contribuindo. Isso resulta em seis itens, os quais, junto com os oito das divisões desiguais do conjunto de quatro membros, equivalem a um total de quatorze itens com contribuição do conjunto de quatro membros. Relativo a apenas esse tipo de comportamento e apenas esses quatro elementos, nós já temos cinquenta itens na lista ordenada associada ao Passo 3 da PCC. Agora siga esse procedimento - em princípio! - para cada agrupamento de todos os mecanismo que contribuem causalmente relativo a cada forma de comportamento inteligente que foi

exibida pelo sujeito em questão (desde que o sujeito tenha exibido uma razoavelmente ampla gama de formas de comportamento inteligente - se não, tudo fica em aberto, pois a riqueza de repertório é uma das características centrais do *explananda* das Ciências Cognitivas).

Com respeito ao exemplo sendo tratado, cada um dos quatro mecanismos presumivelmente aparecerá em muitos subconjuntos com probabilidades condicionais altas (no sentido de que a probabilidade de um subconjunto de um conjunto de contribuintes será alta dado que o complemento do conjunto tem contribuição causal). Essa é uma característica dos mecanismos e da forma de comportamento escolhida. Por exemplo, alguém pode razoavelmente pensar que a probabilidade de que o mecanismo de detecção de bordas contribua dado que o mecanismo que calcula o formato a partir da detecção de sombras está contribuindo seja perto de um; assim, pareceria que todas as vezes que o mecanismo de formato a partir da sombra contribui para o desvio de obstáculos, o mecanismo de detecção de bordas também contribui, ao menos para o sujeito típico, parcialmente porque, como podemos dizer informalmente, eles são ambos mecanismos fundamentais do processamento visual. Similarmente para $P(\text{detecção de bordas} | \text{formato a partir da sombras} \ \& \ \text{distância a partir de disparidade da retina})$ e para $P(\text{distância a partir de disparidade da retina} \ \& \ \text{detecção de bordas} | \text{formato a partir da sombras})$. Note, no entanto, que conjuntos que somente incluem os três mecanismos visuais podem apresentar probabilidades condicionais mais altas que conjuntos que integram o mecanismo de controle motor com os mecanismos visuais, particularmente quando o mecanismo de controle motor estiver sendo condicionado. Parece altamente provável que se os mecanismos visuais estão guiando o desvio do obstáculo, então o mecanismo de controle motor também está. Mas, talvez, o mecanismo de controle motor também contribua para o desvio de obstáculos em casos

nos quais, por exemplo, alguém atravessa com sucesso um cômodo familiar no escuro, com pouca orientação visual. Portanto, $P(\text{formato a partir da sombras} \mid \text{controle motor})$ pode ser significativamente mais baixo que as probabilidades condicionais antes consideradas. Provavelmente não será esse o caso quando o mecanismo de controle motor for condicionado em conjunto com um mecanismo visual. Por exemplo, $P(\text{formato a partir da sombras} \mid \text{controle motor} \ \& \ \text{detecção de bordas})$ provavelmente não será menor que as probabilidades condicionais que envolvem apenas os três mecanismos visuais; pois, se o mecanismo de controle motor em questão contribui em conjunto com o mecanismo de detecção de bordas para desviar do obstáculo, então estaremos quase certamente falando sobre um desvio de obstáculo orientado visualmente, caso no qual é quase certo que o mecanismo de formato a partir da sombras esteja contribuindo também. Como resultado, a consideração dos quatro mecanismos em conexão com desvios de obstáculos presumivelmente resulta em muitos subconjuntos com probabilidades condicionais altas (aquelas que aparecem acima do ponto de corte do Passo 4 da PCC), mesmo se os mecanismos de controle motor apareçam em menos deles do que os outros três.

A PCC foi inicialmente formulada na tentativa de avaliar afirmações sobre cognição estendida, em particular, a afirmação de que as Ciências Cognitivas contemporâneas revelaram que a cognição humana é estendida de uma forma profunda e teoricamente importante (Clark; Chalmers, 1998). Se a cognição deve ocorrer dentro de um sistema cognitivo, como definido pela PCC, então pareceria que, para muitos sujeitos humanos individuais, na maior parte do tempo, o processamento cognitivo ocorre dentro dos limites do corpo do sujeito; pois, em geral, a caracterização do sistema cognitivo que acabamos de dar se opõe a inclusão de ferramentas com finalidades específicas e de uso esporádico,

que tende a ser a condição dos contribuintes causais que estão além dos limites do corpo. (Uma ferramenta com finalidade específica provavelmente aparecerá em muitos conjuntos com probabilidades condicionais altas relativas a uma única forma de comportamento inteligente, mas não aparecerá nos conjuntos relativos a outras formas de comportamento inteligente, colocando esse mecanismo de finalidade específica em uma significativa desvantagem no Passo 5 em comparação com mecanismos que contribuem para uma variedade de formas de comportamento inteligente). No entanto, a localização da cognição humana individual é, em grande medida, uma questão empírica. A proposta baseada em sistemas da PCC deixa aberta a possibilidade de que uma ferramenta - talvez um iPhone (Chalmers, 2008) - que contribui consistentemente para a produção de uma variedade de formas de comportamentos inteligentes através de uma variedade de contextos ao lado de um conjunto variável de co-colaboradores que tenham características similares seja parte do sistema cognitivo de um humano.

Mas, por que pensar que a PCC está correta? Flexibilidade é o centro da cognição e da inteligência - flexibilidade no aprendizado, na aquisição de conceitos e competências, na resolução de problemas e utilização de uma variedade de recursos na perseguição e revisão de objetivos em um ambiente em constante mudança. É essa flexibilidade que atrai a atenção para certas formas de comportamento e performance humanos e motiva o desenvolvimento de uma ciência distintiva (Ciências Cognitivas) para estudá-las, em contraste com tropos ou outras formas de comportamento estereotipadas. É a falta dessa flexibilidade que impulsiona as críticas constantes sobre as formas existentes de Inteligência Artificial. “Não é inteligência de forma alguma” alguém estaria tentado a dizer sobre tais sistemas, “ele não teria qualquer ideia do que fazer se uma situação inesperada surgisse! Ele faz apenas uma coisa!” - seja essa única coisa

jogar xadrez, responder questões de programas de perguntas (*quiz-show*) ou controlar um automóvel.

A PCC está baseada na ideia de que flexibilidade é alcançada em humanos somente pela presença de muitas unidades dispostas de modo a trabalharem juntas em várias combinações. Há abundante evidência de que esse tipo de coisa acontece no cérebro humano (Anderson, 2010, 2014; Cole et al., 2013; Botvinick; Cohen, 2014). Em algumas explicações deste tipo de processo, sub-redes com sobreposição de membros lutam umas com as outras pelo controle por meio de processamentos competitivos. Quando duas sub-redes funcionais têm membros sobrepostos, pode ser necessário apenas um pouco de estímulo diferencial para mudar a atividade do agente da execução de uma tarefa para a execução de outra. Nessa perspectiva, uma mudança de tarefa não requer que uma rede completamente nova assuma o controle de outra rede previamente dominante; mudanças mais sutis na co-ativação de elementos, alguns dos quais já ativos, podem realizar essa transição de modo mais suave. A proposta baseada em sistemas da PCC enfatiza aquilo que parece ser provavelmente a marca central de tal sistema - que qualquer mecanismo seja capaz de cooperar com vários outros subconjuntos de mecanismos para concluir uma variedade de tarefas.

5. Estados cognitivos de grupo?

Agora temos em mãos uma caracterização abstrata do sistema cognitivo que apresenta de forma esquemática um sujeito de estados cognitivos. Dado o quão pouco detalhado é esse sujeito, os proponentes da cognição de grupo podem se sentir confiantes. Parece que o sistema cognitivo de um humano individual - independente de se ele é completamente interno - é pouco mais do que uma coleção

conjuntamente empacotada de “demônios”, todos eles fazendo suas partes individuais para produzir comportamento flexível na ausência de um teatro cartesiano e com pouca coisa em termos de um controlador central (Dennett, 1991). Certamente, em muitos grupos, nós encontramos algo parecido como isto, talvez até mesmo uma coleção de mecanismos relativamente persistentes que contribuem em subconjuntos sobrepostos para a produção de uma ampla variedade de formas de comportamento! Portanto, alguns, talvez muitos, grupos parecem satisfazer as condições baseadas em sistemas necessárias para a posse de estados cognitivos.

Entretanto, dadas as grandes diferenças nos detalhes entre os casos de grupo e individual, devemos saber exatamente o que procurar, se quisermos encontrar cognição no nível do grupo. O material das Seções 3 e 4 sugere o seguinte cenário. Imagine que cientistas sociais completaram estudos muito cuidadosos de estruturas organizacionais, dinâmicas de grupo, etc. De fato, recursos suficientes foram destinados para tais estudos para proporcionar intervenções experimentais. E, imagine que nossos melhores modelos dos dados coletados mostram uma distinção consistente, no nível da arquitetura ou organização de grupo, entre recursos relativamente persistentes e o desfile passageiro de outros recursos que contribuem causalmente para a produção dos fenômenos de interesse. Mas, imagine que quando alguém modele as relações de co-contribuição entre os componentes do pacote persistente, a relação que emerge não é bem capturada pelas probabilidades condicionais das co-contribuições da forma como foi construído na Seção 4. A medida probabilística, cuja adequação melhor captura as relações entre aquilo que aparenta ser as partes obrigatórias do sistema de grupo é, em vez disso, baseada na relação que *altera a probabilidade de co-contribuição de outras duas* (cf. Klein, 2010): tome

quaisquer três mecanismos que contribuem causalmente e pergunte em que medida a contribuição de qualquer um dos três altera a probabilidade da contribuição dos outros dois; então, agrupe os valores resultantes de maneira análoga àquela na qual as probabilidades condicionais das co-contribuições foram agrupadas acima.

Um modelo desse tipo mantém uma relação de ajuste e extensão com o modelo de um sistema cognitivo da Seção 4. Por exemplo, “1. Para qualquer sujeito, em qualquer tempo, modele cada subconjunto não unitário dos mecanismos que possuem contribuição causal distintiva para a produção de qualquer forma de comportamento inteligente...” se tornaria “1. Para qualquer sujeito, em qualquer tempo, modele cada trio de mecanismos que possuem contribuição causal distintiva para a produção de qualquer forma de comportamento inteligente...”, o que claramente constitui um ajuste. Esses modelos compartilham alguma relação probabilística de co-contribuição que liga alguns dos mecanismos que contribuem causalmente em um sistema único, em contraste com a condição dos outros mecanismos contribuintes. Isso fornece um sentido concreto do que seria para um grupo se envolver em um tipo diferente de cognição daquela de um humano individual, mas, ainda assim, cognição.

6. Complicações

Alguns leitores podem se perguntar se eu facilitei demais as coisas para o defensor da cognição de grupo. Em virtualmente qualquer caso no qual um sistema persistente contribui para a produção de algum fenômeno de interesse (em conjunto com vários outros contribuintes causais), é quase certo que *alguma* medida probabilística ou outra faça a partição das causas em dois grupos, onde os membros de um grupo são

componentes do sistema persistente em questão e os membros do outro grupo são as causas restantes. E, pareceria que uma expressão de praticamente qualquer medida probabilística deste tipo pode ser construída pelo ajuste da PCC. Mas, o que torna um modelo substantivo e interessante, alguém poderia pensar, são as interrelações entre os elementos especificados pelo modelo, não meramente por tal fato tão insignificante como o de que os elementos causalmente contribuintes do modelo se dividem em dois grupos.¹⁸ Um modelo conexionista deve incluir unidades de processamento que têm valores de ativação e que podem passar a ativação para outras unidades com as quais elas estão diretamente conectadas. Isso coloca uma limitação significativa na relação “ser o mesmo tipo genérico que” no caso de *ser uma rede conexionista*. Mas, um modelo que meramente aplica medidas probabilísticas para agrupar algumas das causas contribuintes faz poucas demandas sobre um sistema - muito poucas, alguém poderia pensar, para se tomar seriamente como uma teoria da *cognição*, do tipo explanatório-causal.¹⁹ O restante desta seção explora esse problema potencial e outros

¹⁸ Compare a busca por uma concepção genérica de cognição com a busca por uma relação genérica de fundação (*grounding*). Em particular, considere a resposta de Jonathan Schaffer à crítica de Jessica Wilson sobre a tese de que há uma concepção genérica e interessante de fundação (Wilson, 2016). Schaffer argumenta que o que há de comum em várias relações específicas de dependências (as quais são, para Wilson, objetos de interesse genuínos) é que elas podem todas ser representadas usando modelos de equações estruturais (Schaffer, 2016) e que isso produz um entendimento geral útil da relação de fundação. No entanto, alguém poderia razoavelmente questionar se essa propriedade fina de *ser representado por um modelo de equações estruturais* desempenha trabalho explanatório significativo, sem o complemento das dinâmicas relativas a uma relação, domínio ou classe de sistemas particulares. Não está claro, por exemplo, qual conjunto de inferências contrafactuais interessantes estão licenciadas apenas pelo fato de uma relação instanciar a propriedade de *ser representada por um modelo de equações estruturais*.

¹⁹ Considere um ponto correspondente sobre sentenças Ramsey: alguém poderia ficar preocupado que uma sentença Ramsey poderia ser usada para representar uma teoria de qualquer domínio que contenha propriedades interrelacionadas, sem violar a teoria. Seria, portanto, imprudente para um funcionalista em Filosofia da Mente asserir que qualquer sistema que possa ser modelado com o uso de sentenças Ramsey tenha uma mente! (A flexibilidade do formalismo das sentenças Ramsey não deveria ser surpreendente. Considere o uso que Lewis faz dele como uma ferramenta geral para a caracterização de termos teóricos - Lewis, 1970).

dois, concluindo com uma avaliação pessimista a respeito das perspectivas para um realismo sobre estados cognitivos de grupo.

6.1 O problema das condições ausentes

Algo está claramente ausente na caracterização de um sistema cognitivo da Seção 4, ao menos se ela está tentando capturar o tipo *ser cognitivo*. Pense no problema dessa forma: minha caracterização requer, dos mecanismos em questão, que eles contribuam causalmente para a produção de comportamento inteligente; entretanto, a integridade sistêmica em questão parece não ter nada a ver especificamente com cognição. A medida probabilística dos agrupamentos de mecanismos contribuintes poderia ser usada para capturar, por exemplo, sistemas biológicos ou químicos. É em parte por conta dessa preocupação que eu estou inclinado a tratar a PCC como uma articulação de apenas uma condição necessária para que um estado seja cognitivo.²⁰

O que deve ser acrescentado? Uma sugestão óbvia é que o sistema trabalha representacionalmente (Rupert, 2005), talvez representações que são governadas por um tipo específico de algoritmo ou processo (Rupert, 2018b). Ou, alguém pode plausivelmente requerer que para que algo seja um sistema *cognitivo*, ele deva ter estados cognitivos de tipos específicos - isto é, deve, por exemplo, ter algumas memórias, crenças ou estados perceptuais genuínos. Mas, seja lá o que for demandado de uma família de modelos que capturam cognição, nós deveríamos nos preocupar de que a inclusão de tais condições adicionais eleva o nível de exigência a um patamar que grupos típicos não vão satisfazer.

²⁰ Note que tratar isso como uma condição necessária serve suficientemente bem para certos propósitos dialéticos. Se, por exemplo, (a) qualquer lugar em que estados mentais apareçam, deve haver também cognição, (b) a ocorrência de um estado dentro de um sistema cognitivo do tipo caracterizado na Seção 4 é uma condição necessária para que o estado seja cognitivo, e (c) grupos não têm sistemas cognitivos, uma conclusão interessante se segue: grupos não têm estados mentais.

Eis outra maneira de abordar essa preocupação. Em grande medida, uma atitude ecumênica diante da variedade de formas de modelagem científico-cognitivas bem sucedidas motivaram a formulação dos critérios baseados em sistemas da Seção 4. Parte da ideia era mostrar, sem petição de princípio contra, digamos, anti-representacionistas, que uma condição necessária plausível sobre a cognição - uma condição com a qual todos deveriam concordar - deveria ser significativa no debate sobre cognição estendida. Em outros momentos o motivo foi pluralista: talvez todas as formas de modelagem nas Ciências Cognitivas moderadamente bem sucedidas tiveram acesso a alguma coisa importante e, portanto, o que é comum aos programas ilumina a natureza da cognição. O valor dessa postura de mente aberta é devidamente reconhecido, é plenamente possível que uma certa família de modelos da cognição humana vencerá de forma decisiva, na medida em que as Ciências Cognitivas amadureçam; caso no qual as motivações ecumenicas e pluralistas parecerão excessivamente prudentes. Em vez disso, as Ciências Cognitivas terão identificado a cognição, no caso humano, com algum tipo específico de arquitetura e processos, deixando poucas razões para se pensar que o estudo dos nossos casos paradigmáticos de cognição produzirão uma propriedade abrangente de *ser um sistema cognitivo* que grupos possam instanciar - ao invés de considerá-los apenas como sistemas persistentes, o que muitos grupos claramente são.

6.2 O problema do comportamento inteligente e flexível

Muitos grupos simplesmente não produzem uma ampla gama de formas de comportamento inteligente. Aqueles que aparecem frequentemente na literatura sobre cognição de grupo exibem um alto grau de especificidade de tarefas: a Suprema Corte dos EUA produz

decisões, o mercado de ações prediz a performance futura de empresas, um time de basquete joga basquete, um grupo de marinheiros navega, mas em cada caso o grupo não faz muitas outras coisas, enquanto um grupo coeso.

Por outro lado, sistemas cognitivos apropriados produzem uma ampla gama de formas de comportamento inteligente; a caracterização de qualquer forma de comportamento de um sujeito como sendo inteligente depende da produção de uma ampla gama de outras formas de comportamento inteligente pelo sujeito. A capacidade de um sujeito singular de produzir uma ampla gama de formas de comportamento de maneira flexível constitui um dos *explananda* centrais das Ciências Cognitivas. Como notado acima, uma crítica constante sobre a inteligência artificial se volta exatamente para esse ponto: as façanhas do Deep Blue, MYCIN, carros autônomos, Alpha Go e muitos outros sistemas especializados são muito restritas para a manifestação de uma inteligência genuína. Inteligência requer a capacidade de resolver todo tipo de problema flexivelmente - priorizar tarefas, alternar entre tarefas de forma apropriada em um momento apropriado, alternar entre estratégias para o cumprimento de uma determinada tarefa conforme a situação exigir, e assim por diante (contar uma piada ou cantar uma música quando for necessário, estudar para o teste de admissão da Faculdade de Medicina, quando isso for necessário, etc.).

Grupos típicos não parecem ter a estrutura certa para exibir uma ampla gama de formas de comportamento inteligente. E, na medida em que grupos exibem alguma flexibilidade, uma certa assimetria na produção de tais comportamentos parece reveladora. É praticamente sempre o caso que onde um grupo exibe algum grau de flexibilidade y , cada indivíduo contribuinte no grupo que ajuda na produção do comportamento em questão é capaz de flexibilidade cognitiva de grau x ,

onde x é significativamente maior do que y . Isso parece ser o oposto nas situações dos casos de humanos individuais. Humanos individuais podem até ter subsistemas especializados responsáveis pela execução de grande parte do trabalho que sejam de suas alçadas. Entretanto, o sujeito individual como um todo tem uma capacidade muito mais flexível do que qualquer módulo individual ou subsistema especializado. Ao contrário, ações de grupo são praticamente sempre levadas a cabo, ou implementadas, pelos membros individuais (ou representantes) do grupo, membros estes que são individualmente muito mais flexíveis do que o grupo. Talvez esse fato, por si só, não deponha contra o realismo sobre estados de grupo. Mas, essa assimetria - o fato de que os indivíduos são mais flexíveis do que o grupo - deveria nos fazer hesitar; em tais casos, parece que seres genuinamente inteligentes estão cooperando para organizar um grupo que é, como no caso da I.A., somente uma réplica relativamente inflexível de uma habilidade especializada.²¹

6.3 O problema das restrições na seleção de modelos

Nós apenas derivamos conclusões substantivas das relações de ajuste e extensão quando os modelos em questão são nossos melhores modelos para os fenômenos relevantes. Portanto, cânones de seleção de modelos - e considerações metodológicas, falando de maneira mais geral - incidem no debate sobre cognição de grupo. Isso inclui desideratos metodológicos tais como simplicidade, conservadorismo e a unidade da ciência, bem como os princípios orientadores da modelagem estatística.

A contribuição dos princípios metodológicos ameaça causar problemas para os realistas. Se, por exemplo, uma restrição para que

²¹ Agradeço a Gabriel Rabin por ter me incentivado a esclarecer minha opinião sobre essas questões.

estados sejam cognitivos é de que eles sejam representacionais, então questões sobre, por exemplo, parcimônia quantitativa (Nolan, 1997) mostram as caras. Considere um caso no qual todos os membros de um tribunal de apelação votem para manter uma condenação criminal, mesmo que nenhum dos membros acredite que a parte em questão seja culpada. Isso mostra que o tribunal está em um estado representacional - talvez uma crença de que o acusado seja culpado - que difere, e existe além, dos estados cognitivos dos membros individuais do tribunal? Considerações de parcimônia quantitativa sugerem que não. Postular um estado representacional adicional - o estado cognitivo do tribunal - é gratuito e, portanto, não se enquadraria no modelo do fenômeno em questão (isto é, a emissão de uma decisão por parte do tribunal). Pois, presumivelmente, cada membro individual do tribunal tem suas razões para votar pela manutenção da condenação, mesmo que não acredite que o acusado seja culpado. Tais razões podem variar desde uma adesão ao princípio dos precedentes, uma crença de que haverá uma revolta se o acusado for libertado, até uma crença de que decisões judiciais devem adotar um modelo baseado em premissas. Entretanto, uma vez que essas razões tenham sido incluídas no modelo de emissão de decisões pelo tribunal, todos os recursos representacionais necessários para explicar o evento estão à mão (Rupert, 2005, 2011, 2014; Wilson, 2004). Adicionar mais do mesmo, isto é, mais recursos representacionais, é gratuito. *Ceteris paribus*, portanto, cânones de seleção de modelos impõem a escolha de um modelo no qual os estados representacionais não aparecem como estados do grupo como um todo. Desta forma, nosso melhor modelo de processos envolvendo grupos que produzem comportamento inteligente (se tal comportamento existir) provavelmente não atribuirá estados representacionais para grupos, contra a Premissa 1 de List.

Essa moral pode ser expressa não tanto em termos de parcimônia, mas como uma questão sobre redução e a unidade da ciência. É um cânone da seleção de modelos que, outras coisas permanecendo iguais, nós devemos preferir modelos do fenômeno que se encaixam com modelos bem sucedidos em domínios adjacentes. Ou, atenuando um pouco, é um cânone da seleção de modelos que, *ceteris paribus*, quando estivermos tentando derivar conclusões ontológicas a partir de modelos bem sucedidos, devemos favorecer as coleções de modelos que apresentem uma imagem unificada da natureza sobre as coleções que não o fazem. Considere, por um lado, um modelo de acordo com o qual instituições políticas e sociais têm estados cognitivos, porque elas exibem racionalidade ao longo do tempo, e, por outro lado, um modelo de acordo com o qual somente processos cognitivos de humanos individuais expliquem fenômenos mensuráveis relacionados às instituições políticas e sociais (ignorando todos os outros componentes de ambos os tipos de modelos, tais como folhas de papel, bombas de gasolina, etc.). A última proposta é mais consistente com a ênfase na unidade da ciência. Ou, alguém poderia combinar essas duas questões na forma de um dilema: ou o proponente de estados cognitivos de grupos opta por processos de grupo misteriosos e independentes, caso onde afronta a unidade da ciência, ou faz um esforço genuíno de integrar modelos em um único pacote e, portanto, afronta a parcimônia quantitativa, se insistir que grupos têm estados cognitivos próprios.²²

²² Poderia ser objetado que recorrer aos cânones de seleção de teorias é muito severo, pois o status desses cânones é contestado e sua contribuição histórica para o pensamento científico é complexo e opaco. Poderia ser pensado que, ao contrário, melhor seria limitar a discussão a considerações de razões de primeira ordem a favor ou contra as explicações oferecidas para fenômenos, dados ou efeitos específicos. Acato esse ponto. Mas, até onde posso ver, as preocupações expressas aqui podem ser traduzidas em argumentos de primeira ordem caso a caso: nós sabemos por que os juízes votaram da forma como o fizeram e, portanto, nós sabemos por que a decisão tomada foi aquela, sem qualquer referência a estados cognitivos de grupo. E essa maneira de ver a questão dispensa qualquer

Entretanto, é preciso ter cuidado para não se deixar enganar pela conversa sobre redução. Voltando ao argumento realista de List apresentado na Seção 2, que sugere vagamente uma linha de pensamento anti-reducionista - a ideia sendo que se as Ciências Sociais fazem um uso produtivo das atribuições de crenças e desejos de grupo, um naturalista de carteirinha não deveria resistir (a menos que os estados em questão pudessem ser reduzidos). Essa vaga impressão é reforçada por várias passagens em List e Pettit (2011) sobre, por exemplo, relações entre níveis (p. 4), competição causal multinível (p. 161), a realização de atitudes de grupos por atitudes individuais (p. 77), superveniência (como uma relação entre níveis; p. 65) e as dificuldades em reduzir atitudes de grupo a atitudes individuais (p. 76-77, 194). Entretanto, esse enquadramento da questão me parece confuso ou, ao menos, enganador. De forma geral, a questão da redução é uma questão *entre* níveis (como são as questões sobre superveniência e realização), que surgem em casos de diferentes domínios científicos com diferentes conjuntos de propriedades e isso é geralmente colocado em termos de relações entre diferentes níveis. No entanto, esse, decididamente, não é o tipo de caso que temos em mãos. Os estados cognitivos de grupos e individuais são supostamente do mesmo tipo, partes do mesmo domínio: o mental, psicológico ou cognitivo. Isso é o que deveria ser chocante sobre o realismo em questão. Afinal de contas, não seria surpreendente descobrir que algum estado ou outro do nível de grupo desempenha um papel explanatório-causal (cf. “o peso do grupo causou o colapso do terraço”). List e Pettit parecem perder isso de vista quando fazem aquilo que consideram como uma comparação a favor do

necessidade de apelar para um princípio metodológico abrangente. Agradeço Peter Achinstein por me pressionar sobre esse ponto.

realismo entre (a) a relação entre estados cognitivos de grupos e estados cognitivos individuais e (b) a relação entre estados cognitivos individuais e “a configuração e funcionamento de subsistemas biológicos” (p. 7); eles argumentam que, se é legítimo resistir a redução no caso (b), então o caso (a) deveria ser tratado do mesmo modo (p. 78; cf. Goldstone; Gureckis, 2009; Huebner, 2008). Entretanto, esse argumento por analogia é falho. O caso (b) trata dois conjuntos genuínos de propriedades diferentes (propriedades cognitivas, de um lado, e propriedades biológicas, de outro), adequadamente levantando questões sobre redução e anti-redução. Mas, o caso (a) diz respeito ao mesmo tipo de estados e propriedades, em ambos os lados da comparação, não um conjunto de propriedades ou estados de um tipo realizando ou sobrevivendo a propriedades e estados de outro tipo.²³ Portanto, alguém não poderia simplesmente assumir os recursos anti-reducionistas utilizados no caso (b) para neutralizar argumentos baseados em parcimônia dirigidos ao caso (a).²⁴

Aqui há outra forma de ver esse ponto: no caso (b), redução, no sentido inter-teórico, é bloqueada por argumentos anti-reducionistas (Fodor, 1974, por exemplo) e, dessa forma, certificamos ontologicamente um domínio de propriedades - psicológicas, mentais ou cognitivas.

²³ List e Pettit algumas vezes parecem reconhecer esse fato (2011, p. 1, 78), mas não parecem considerar o quanto isso compromete sua defesa de grupos agentes.

²⁴ Em resposta a isso, alguém poderia apelar para a grande aceitação de níveis pessoais e subpessoais (ao invés da relação entre os níveis psicológicos e neurais) nas teorias cognitivas (Huebner, 2008). Se, desconsiderando a força da parcimônia, recusarmos eliminar o nível pessoal em favor do subpessoal, não deveríamos permitir que o nível individual exclua o nível do grupo, segue o raciocínio. Na minha opinião, no entanto, a distinção entre níveis pessoais e subpessoais não desempenha qualquer papel substantivo nas Ciências Cognitivas (veja Rupert, 2018a, para argumentos nesse sentido); caso no qual, o realista não pode encontrar consolo aqui. Além do mais, muitos filósofos que aceitam a existência de um nível pessoal consideram que os estados que ali aparecem são de um tipo diferente daqueles que aparecem no nível subpessoal. Portanto, a suposta distinção entre níveis pessoais e subpessoais não nos fornece um exemplo de um caso no qual as mesmas propriedades psicológicas, mentais ou cognitivas apareçam em diferentes níveis.

Entretanto, o caso (a) envolve uma situação na qual nós já permitimos a entrada de tipos mentais ou cognitivos em nossa ontologia e a única questão é até onde nós os aplicamos. O caso (b) é uma questão entre níveis; o caso (a) é uma questão dentro de um nível. Essa é a razão de porque o argumento por analogia de List e Pettit é inadequado e no caso (a), no caso que temos em mãos, deveríamos focar na parcimônia quantitativa (Nolan, 1997), a questão de quantas instâncias de um dado tipo de estado, já aprovado, ou propriedade deveríamos incluir em nossos modelos científicos.²⁵

Considere uma resposta final para a questão sobre parcimônia e redução. List e Pettit (2011) afirmam:

A agência dos grupos se relaciona de uma forma tão complexa com a agência de indivíduos que nós temos poucas chances de rastrear as disposições de um grupo agente e interagir com ele enquanto um agente para contestar ou interrogar, persuadir ou coagir, se nós conceituarmos suas ações no nível individual. (p. 76)

Eu acho esse comentário desconcertante. Ao longo de seu livro, List e Pettit desenvolvem análises precisas de como melhor podemos construir grupos para que os grupos possam exibir racionalidade (da

²⁵ Em uma discussão de meus trabalhos prévios sobre cognição de grupo (Rupert, 2014), Himmelreich (2015) reclamou que, ao oferecer um argumento baseado na parcimônia quantitativa contra o realismo, eu não utilizei várias ferramentas conceituais produzidas nos debates sobre exclusão causal. Entretanto, dado os comentários no corpo do texto, pode parecer claro porque o caminho proposto por Himmelreich é menos promissor do que ele considera. As ferramentas conceituais desenvolvidas nos debates sobre exclusão causal são, geralmente, direcionadas para relações entre níveis; no diagrama padrão de Kim (e.g., Kim, 1998), as relações verticais são relações de superveniência, as quais são, normalmente, tomadas como sendo instanciadas entre famílias de propriedades distintas. Mas, a relação entre estados de grupos e estados individuais não é entendida pelo realista como dizendo respeito a relações entre conjuntos de propriedades distintas. Afinal de contas, se o conjunto de propriedades for distinto, porque o realista usa os mesmos termos - e.g., 'crença' e 'desejo' - para falar sobre estados de grupo como são usados para falar sobre estados individuais? Note, também, que quando List e Pettit (2011) seguem o caminho proposto por Himmelreich, eles abordam o problema especificamente como um problema sobre causalção multinível (p. 161), o que não se adequa com o caso de estados cognitivos competindo com estados cognitivos; isso não seria uma questão de causalção multinível, pelas razões dadas acima.

melhor forma possível, considerando os vários resultados de impossibilidade). Além disso, essas formulações nos dizem exatamente com quem falar em uma organização, porque elas nos dizem, de forma relativamente transparente, quem estará votando (ou, alternativamente, tendo a palavra decisiva) sobre o que. Por exemplo, List e Pettit demonstram formalmente que, sob certas condições, um procedimento efetivo para a manutenção da racionalidade de grupo é dividir questões logicamente interrelacionadas em uma agenda da seguinte forma: identifique as premissas, atribua um subgrupo de membros com conhecimento especializado relevante para cada uma das premissas, adote uma votação baseada na regra da maioria em cada subgrupo a respeito da premissa a ele atribuída e, então, chegue a uma conclusão com base somente nos resultados desses votos (isto é, siga um procedimento baseado nas premissas) (List; Pettit, 2011, p. 95). Neste caso, parece claro a quem alguém deveria se dirigir se quisesse persuadir o grupo a chegar à sua conclusão preferida ou para forçar um determinado resultado: os membros dos subgrupos especializados que julgar que tenha mais chances de convencer. (Existe um setor inteiro nos Estados Unidos - o setor do *lobby* - baseado nesse tipo de observação). O que List e Pettit tomam como sendo atitudes de grupo sobrevivem de atitudes individuais e, como eles enfatizam, algumas vezes isso não se dá de uma maneira direta. Isso é razoável. Mas, a maior parte do livro deles demonstra exatamente por que e por quais razões o padrão de superveniência deveria ser de uma maneira não inteiramente direta (*less-than-entirely-straightforward*). E, ao explicar essas relações não inteiramente diretas com decisões individuais que fixam fatos supervenientes sobre o grupo, List e Pettit demonstram que uma relação de superveniência não inteiramente direta não implica o tipo de opacidade que poderia motivar a postulação de estados autônomos de

grupos. Eles fornecem formulações para a construção de estados de grupo por indivíduos e, portanto, fornecem um guia para conceituar esses estados como amálgamas de decisões individuais (cf. Rupert, 2005, p. 179; Roth, 2014)²⁶ - não enquanto estados genuínos de grupos - o que nos diz quais indivíduos deveriam ser influenciados sobre quais questões em qual estágio de um determinado processo para conduzir à um resultado que gostaríamos.

Conclusão

Nas Seções 2-5, eu defendi uma proposta de teoria sobre tipos naturais enquanto uma ferramenta para avaliação do realismo sobre cognição de grupo; ofereci uma teoria do tipo de evidência pertinente para a determinação da igualdade de tipo natural; e apliquei a teoria para o caso de um tipo de *cognição* abrangente, na esperança de oferecer ao realista um caminho a seguir. Na seção final, dificuldades proibitivas foram apresentadas. Eu concluí, então, que as perspectivas para a descoberta de estados cognitivos genuínos em grupos parecem fracas, embora muito dependa de trabalho empírico ainda a ser feito.

Particularmente, ainda que essa seção final tenha sido em grande parte pessimista na sua avaliação das perspectivas para o realismo, é

²⁶ Rupert (2005) e Roth (2014) enfatizam o papel de estados individuais de ordem superior nas críticas ao realismo baseadas em redução ou parcimônia, tais estados de ordem superior são tidos como endossos individuais de regras de agregação ou corroboração a procedimentos canônicos para a construção de uma conclusão oficial declarada pelo grupo. Um exemplo desse tipo de procedimento canônico no campo jurídico é o de seguir os precedentes. É uma coisa acreditar que a decisão legal correta é P, baseado somente na forma como a lei foi escrita. É outra coisa acreditar que o julgamento correto é não-P, baseado somente na forma como a lei foi escrita, mas que a melhor coisa para o tribunal fazer no balanço final, dados os precedentes relacionados à P, é decidir que P. Um juiz com um comprometimento de ordem superior com o papel dos precedentes, pode muito bem estar na última situação. Ao defender o realismo contra as críticas baseadas em redução ou parcimônia, List e Pettit, assim como Gilbert, argumentam como se o crítico tenha apenas estados de primeira ordem dos indivíduos (tais como a crença de que o réu é culpado) para trabalhar. Isso, no entanto, é argumentar contra um espantinho do individualismo.

válido mencionar uma empreitada empírica digna de ser pesquisada. Pode ser que, mesmo que a PCC não pareça capturar uma propriedade robusta ou interessante - para além de algo ser um sistema coerente - o tipo de estrutura caracterizada é, de fato, exatamente o tipo de estrutura necessária para sustentar comportamento flexível (alternância oportuna entre tarefas, etc.). Talvez, falando nomologicamente, não há outra forma de construir um sistema que exiba flexibilidade do tipo certo, exceto dotando-a com uma ampla coleção de mecanismos que contribuam em subconjuntos sobrepostos para a produção de uma ampla gama de formas de comportamento; talvez seja essa estrutura em particular que torne possível a flexibilidade. Nesse caso, tal estrutura teria mais a ver com a natureza da cognição do que algumas vezes sugerido acima e a busca por tal estrutura em grupos pode, no fim das contas, dar frutos relacionados à cognição.

Referências

- ANDERSON, M. L. Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, 33(4), 245–313, 2010.
- ANDERSON, M. L. *After Phrenology: Neural Reuse and the Interactive Brain*. Cambridge, MA: MIT Press, 2014.
- ARICO, A. Folk psychology, consciousness, and context effects. *Review of Philosophy and Psychology*, 1(3), 371–393, 2010.
- BARNIER, A. J.; SUTTON, J.; HARRIS, C. B.; WILSON, R. A. A conceptual and empirical framework for the social distribution of cognition: The case of memory. *Cognitive Systems Research*, 9(1), 33–51, 2008.
- BLOCK, N. Troubles with functionalism. In C. SAVAGE (Ed.), *Perception and cognition. Issues in the foundations of psychology*, Minnesota studies in the philosophy of science (Vol. 9). Minneapolis: University of Minnesota Press, p. 261-325, 1978.

- BLOCK, N. (1997). Anti-reductionism slaps back. In J. TOMBERLIN (Ed.), *Mind, Causation, World: Philosophical Perspectives* 11 (pp. 107–133). Boston, MA: Blackwell.
- BOTVINICK, M. M.; COHEN, J. D. The computational and neural basis of cognitive control: Charted territory and new frontiers. *Cognitive Science*, 38(6), 1249–1285, 2014.
- BROOKS, R. A. *Cambrian intelligence: The early history of the new AI*. Cambridge, MA: MIT Press, 1999.
- CHALMERS, D. Foreword to Andy Clark, *Supersizing the mind*. In A. CLARK. *Supersizing the mind: Embodiment, action, and cognitive extension* (pp. ix–xvi). New York, NY: Oxford University Press, 2008.
- CHURCHLAND, P. M. Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78(2), 67–90, 1981.
- CLARK, A. *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press, 1997.
- CLARK, A. *Supersizing the mind: Embodiment, action, and cognitive extension*. New York, NY: Oxford University Press, 2008.
- CLARK, A. Finding the mind. *Philosophical Studies*, 152(3), 447–461, 2011.
- CLARK, A.; Chalmers, D. The extended mind. *Analysis*, 58(1), 7–19, 1998.
- COLE, M. W.; REYNOLDS, J. R.; POWER, J. D.; REPOVS, G.; ANTICEVIC, A.; BRAVER, T. S. Multitask connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, 16(9), 1348–1355, 2013.
- DE BRIGARD, F. Cognitive systems and the changing brain. *Philosophical Explorations*, 20(2), 224–241, 2017.
- DENNETT, D. *Consciousness explained*. Boston, MA: Little, Brown and Company, 1991.
- EVANS, J. St. B. T.; FRANKISH, K. (Eds.). *In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press, 2009.

FODOR, J. Special sciences (Or: The disunity of science as a working hypothesis). *Synthese*, 28 (2), 97–115, 1974.

FODOR, J. *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press, 1983.

FODOR, J. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press, 1987.

GENDLER, T. Alief and belief. *Journal of Philosophy*, 105(10), 634–663, 2008.

GIERE, R. The role of agency in distributed cognitive systems. *Philosophy of Science*, 73(5), 710–719, 2006.

GILBERT, M. Collective epistemology. *Episteme: A Journal of Social Epistemology*, 1(2), 95–107, 2004.

GOLDSTONE, R.; GURECKIS, T. Collective behavior. *Topics in Cognitive Science*, 1(3), 412–438, 2009.

GOLDSTONE, R.; G. THEINER. The multiple, interacting levels of cognitive systems (MILCS) perspective on group cognition. *Philosophical Psychology* 30 (3), 338–372, 2017.

GRIFFITHS, T. L.; CHATER, N.; KEMP, C.; PERFORNS, A.; TENENBAUM, J. B. Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, 14(8), 357–364, 2010.

GRUSH, R. The architecture of representation. *Philosophical Psychology*, 10(1), 5–23, 1997.

HIMMELREICH, J. Review of From Individual to Collective Intentionality: New Essays, *Economics and Philosophy*, 31(3), 479–486, 2015.

HUEBNER, B. Do you see what we see? An investigation of an argument against collective representation. *Philosophical Psychology*, 21(1), 91–112, 2008.

HUEBNER, B. *Macrocognition: A theory of distributed minds and collective intentionality*. Oxford, UK: Oxford University Press, 2013.

- HUEBNER, B.; BRUNO M.; SARKISSIAN, H. What does the nation of China think about phenomenal states? *Review of Philosophy and Psychology*, 1(2), 225–243, 2010.
- HUTCHINS, E. *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.
- KIM, J. *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge, MA: MIT Press, 1998.
- KLEIN, C. Critical Notice: Cognitive systems and the extended mind by Robert Rupert. *The Journal of Mind and Behavior*, 31(3–4), 253–264, 2010.
- KNOBE, J.; & PRINZ, J. Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences*, 7(1), 67–83, 2008.
- LEWIS, D. How to define theoretical terms. *Journal of Philosophy*, 67(13), 427–446, 1970.
- LEWIS, D. Mad pain and Martian pain. In N. BLOCK (Ed.), *Readings in the Philosophy of Psychology*, Volume one. Cambridge, MA: Harvard University Press, p. 216–222, 1980.
- LIST, C. What is it like to be a group agent? *Noûs*, 52(2), 295–319, 2018.
- LIST, C.; PETTIT, P. *Group agency: The possibility, design, and status of corporate agents*. Oxford, UK: Oxford University Press, 2011.
- MARR, D. *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press, 1982.
- NOLAN, D. Quantitative parsimony. *British Journal for the Philosophy of Science*, 48(3), 329–343, 1997.
- PUTNAM, H. Psychological predicates. In W. H. CAPITAN; D. D. MERRILL (Eds.), *Art, Mind, and Religion*. Pittsburgh, PA: University of Pittsburgh Press, p. 37–48, 1967.
- PUTNAM, H. The meaning of ‘meaning’. In H. PUTNAM (Ed.), *Mind, Language, and Reality: Philosophical Papers* (Vol. 2). Cambridge, UK: Cambridge University Press, p. 215–271, 1975.
- ROTH, A. S. Indispensability, the discursive dilemma, and groups with minds of their own. In S. R. CHANT; F. HINDRIKS; G. PREYER (Eds.) *From Individual to Collective Intentionality: New Essays*. Oxford, UK: Oxford University Press, p. 137–162, p. 2014.

- ROVANE, C. Group agency and individualism. *Erkenntnis*, 79(9), 1663–1684, 2014.
- RUMELHART, D. E.; HINTON, G. E.; MCCLELLAND, J. L. A general framework for parallel distributed processing. In D. E. RUMELHART; J. L. MCCLELLAND; PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, Foundations. Cambridge, MA: MIT Press, p. 45–76, 1986.
- RUPERT, R. D. Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 101(8), 389–428, 2004.
- RUPERT, R. D. Minding one's cognitive systems: When does a group of minds constitute a single cognitive unit? *Episteme: A Journal of Social Epistemology*, 1(3), 177–88, 2005.
- RUPERT, R. D. *Cognitive systems and the extended mind*. Oxford, UK: Oxford University Press, 2009.
- RUPERT, R. D. Extended cognition and the priority of cognitive systems. *Cognitive Systems Research*, 11(4), 343–56, 2010.
- RUPERT, R. D. Empirical arguments for group minds: A critical appraisal. *Philosophy Compass*, 6(9), 630–639, 2011.
- RUPERT, R. D. Memory, natural kinds, and cognitive extension; or, Martians don't remember, and cognitive science is not about cognition. *Review of Philosophy and Psychology*, 4(1), 25–47, 2013.
- RUPERT, R. D. Against group cognitive states. In S. CHANT; F. HINDRIKS; G. PREYER (Eds.), *From individual to collective intentionality*. Oxford, UK: Oxford University Press, p. 97–111, 2014.
- RUPERT, R. D. The self in the age of cognitive science. *Philosophic Exchange*, 47(1), 2018a.
- RUPERT, R. D. Representation and mental representation. *Philosophical Explorations*, 21(2), 204–225, 2018b.
- SCHAFFER, J. Ground rules: Lessons from Wilson. In K. AIZAWA & C. GILLET (Eds.), *Scientific composition and metaphysical ground*. Springer, p. 143–169, 2016.

- SCHWITZGEBEL, E. If materialism is true, the United States is probably conscious. *Philosophical Studies*, 172(7), 1697–1721, 2015.
- SHOEMAKER, S. Some varieties of functionalism. *Philosophical Topics*, 12(1), 93–118, 1981.
- SPREVAK, M. Extended cognition and functionalism. *Journal of Philosophy*, 106(9), 503–27, 2009.
- STICH, S. *Deconstructing the Mind*. Oxford, UK: Oxford University Press, 1996.
- SYTSMA, J.; MACHERY, E. How to study folk intuitions about phenomenal consciousness. *Philosophical Psychology*, 22(1), 21–35, 2009.
- SYTSMA, J.; MACHERY, E.. Two conceptions of subjective experience. *Philosophical Studies*, 151(2), 299–327, 2010.
- THEINER, G.; & O'CONNOR, T. The emergence of group cognition. In A. CORRADINI; T. O'Connor (Eds.). *Emergence in science and philosophy*. New York, NY: Routledge, p. 78–117, 2010.
- THEINER, G.; Allen, C.; GOLDSTONE, R. L. Recognizing group cognition. *Cognitive Systems Research*, 11(4), 378–395, 2010.
- TOLLEFSEN, D. From extended mind to collective mind. *Cognitive Systems Research*, 7(2–3), 140–50, 2006.
- WHEELER, M. In search of clarity about parity. *Philosophical Studies*, 152(3), 417–425, 2011.
- WILSON, J.. The unity and priority arguments for grounding. In K. AIZAWA; C. GILLET (Eds.), *Scientific Composition and Metaphysical Ground* (pp. 171–204). London, UK: Palgrave Macmillan, 2016.
- WILSON, M. Six Views of Embodied Cognition. *Psychonomic Bulletin and Review*, 9(4), 625–36, 2002.
- WILSON, R. A. *Boundaries of the Mind: The Individual in the Fragile Sciences*. Cambridge, UK: Cambridge University Press, 2004.

9

INTROSPECÇÃO EM MENTES DE GRUPOS, DESUNIÃO DA CONSCIÊNCIA E PESSOAS INDISCRETAS ¹

Eric Schwitzgebel

Sophie R. Nelson

Tradução: Anderson Fonseca, Luiz Paulo Cichoski

Ordinariamente, nós pensamos que pessoas são entidades contáveis, discretas. Uma sala de conferências contém zero pessoas, ou uma, ou duas, ou vinte e três, ou algum outro número inteiro. Se você perguntar quantas pessoas estão presentes e alguém responder “duas e três quartos”, você provavelmente pensará que ele está brincando. (Talvez ele esteja contando um feto como uma pessoa parcial). Se ele responder “é indeterminado entre dois e vinte e três”, você provavelmente pensará que a indeterminação diga respeito ao critério de presença (será que participar por meio de uma chamada de vídeo conta como estando “presente”?) ao invés de dizer respeito à discrição² de pessoas.

A discrição entre pessoas é um fato contingente sobre (a maioria?) os seres humanos existentes na Terra, não um fato fundamental sobre a natureza de sistemas cognitivamente sofisticados, como ficará evidente quando considerarmos como a introspecção pode funcionar em entidades de grupo. Uma consequência da indiscrição de pessoas é a indeterminação a respeito da fronteira entre a introspecção dentro de uma mente e a comunicação entre mentes, que implica uma dimensão

¹ Tradução do texto: SCHWITZGEBEL, E.; NELSON, S. Introspection in Group Minds, Disunities of Consciousness, and Indiscrete Persons. *Journal of Consciousness Studies*, v. 30, n. 9-10, p. 188-203, 2023; DOI: <https://doi.org/10.53765/20512201.30.9.188>. Agradecemos aos autores e editores da revista pela autorização da publicação desta tradução. (N.T.)

² O termo “discreto” é usado pelo autor no sentido algébrico, isto é, referindo-se, aqui, a grandezas descontínuas, isto é, constituída por unidades distintas; enquanto o termo “indiscreto” atribui-se a grandezas contínuas, formada por unidades indistintas entre si. (N.T.)

de variedade introspectiva além daquelas descritas por Kammerer e Frankish (2023).

1. Mentes Ancilares: Percepção, Ação e Memória

Vamos considerar um sistema cognitivo com aquilo que vamos chamar de estrutura ancilar, inspirada na arquitetura cognitiva visionada pelo romance de ficção científica de Ann Leckie (2018) *Justiça Ancilar*. Um computador sofisticado orbita um planeta. Na superfície do planeta há dez pelotões de vinte robôs humanoides cada. O computador em órbita e os robôs humanoides estão em constante comunicação via satélite. Sem pressupor qualquer coisa sobre as fronteiras de pessoas, vamos chamar o computador em órbita de Aleph e vamos chamar cada robô humanoide pela letra do pelotão e o número de associado, de A1 até J20. Para os propósitos deste artigo, vamos assumir que cognição e consciência muito parecidas com a humana são possíveis em computadores e que processos funcionais e informacionais são aquilo que importa para a consciência. (Se consciência requer um cérebro, talvez possamos imaginar que os robôs tenham cérebros orgânicos modificados com interfaces computacionais, o que estaria muito mais próximo da proposta original de Leckie).

Se todo o processamento cognitivo ocorre em Aleph, enquanto os robôs são simplesmente dispositivos para entrada de dados que transmitem informações brutas a partir de câmeras e microfones diretamente para Aleph, então o sistema ancilar presumivelmente tem uma mente com muitos sistemas de entrada de dados distribuídos. Se, ao contrário, todo o processamento cognitivo ocorre nos robôs e Aleph é apenas um sistema de transmissão, então o sistema ancilar presumivelmente tem 200 mentes, cada uma em um local distinto da

superfície do planeta. Mais interessantes são os casos intermediários. (Se você estiver se perguntando se poderia haver mentes robóticas individuais, como também uma mente de grupo, segure esse pensamento por enquanto.)

Suponha, então, que processamentos substanciais ocorram em cada robô antes que seu sinal seja transmitido e que processamentos substanciais também ocorram em Aleph depois que os sinais são recebidos, moldando os sinais que Aleph emite de volta aos robôs. Novamente, os casos limites são de interesse secundário. Se os sinais têm relativamente pouco conteúdo e cada um deles precisa ser substancialmente processado antes de moldar a opinião e comportamento do receptor, haverá 201 mentes (Aleph e os 200 robôs, todos seriam indivíduos distintos). Se os sinais são rápidos e robustos, com acesso direto aos centros cognitivos e motores - de forma que seja arbitrário dizer que alguma parte do processamento ocorre em um robô ou em Aleph - então haverá uma única mente distribuída pelas 201 localizações. Esse caso é intermediário: A conectividade entre os centros de cognição dos robôs e Aleph é rápido, robusto e direto, mas não o suficiente para que a entidade ancilar constitua um caso claro de uma única mente unificada.

Para melhor vislumbrar essa possibilidade, considere alguns processos perceptuais, motores e de memória distribuídos.

Como poderia a percepção funcionar? Suponha que o pelotão A esteja examinando um campo de flores. Dezesesseis membros de A, cada um empregando seu próprio processamento local, inicialmente representam as flores como sendo 50% azuis e 50% vermelhas. Quatro membros de A, ao contrário, inicialmente representam a mesma coleção de flores como sendo 70% azuis e 30% vermelhas: Aleph recebe esses sinais conflitantes. Talvez Aleph adote um procedimento *winner-take-*

all^3 e represente as flores como sendo 50% azuis. Talvez Aleph adote um procedimento de média ponderada e represente as flores como sendo $50\% * (16/20) + 70\% * (4/20) = 54\%$ azuis. Talvez Aleph adote graus de crença: um grau de crença 0,8 para a distribuição 50/50 e um grau de crença 0,2 para a distribuição 70/30; ou, alternativamente, um grau de crença 0,7 para 50/50; um grau de crença 0,2 para 60/40 e um 0,1 para 70/30. Certamente, há muitas outras alternativas menos simples.

As representações de Aleph - suponha, para simplificar, que Aleph adote um procedimento *winner-take-all*, representando a proporção como sendo 50/50 - é, então, transmitida aos pelotões. Os pelotões de B até J recebem essa nova informação, nunca tendo visto o campo. Dezesseis membros de A recebem o sinal de Aleph como confirmações de suas representações locais de 50/50. Quatro membros de A recebem o sinal de Aleph como desconfirmações. Eles podem, então, atualizar suas representações locais para 50/50, descartando suas representações iniciais. Alternativamente, eles podem adotar um meio termo entre suas representações iniciais e a representação recebida de Aleph: talvez 60/40 ou talvez um grau de crença 0,5 para 70/30 e um grau de crença 50% para 50/50. Novamente, isso apenas é o início das possibilidades. Mais complexidades podem ser introduzidas se as representações de alguns robôs merecem um peso maior em virtude de eles terem melhores pontos de vista ou melhores equipamentos, ou se os robôs podem comunicar diretamente, sem que Aleph seja um intermediário.

Como as ações poderiam funcionar? Nós podemos imaginar, novamente, casos extremos: ações completamente controladas por Aleph, onde os robôs seriam membros passivos, ou ações

³ O *Winner-take-all* é um princípio computacional para modelos de redes neurais, no qual múltiplos neurônios competem entre si por ativação; porém, apenas aquele com maior nível de estímulo é ativado, inibindo os demais neurônios. (N.T.)

completamente controladas pelos robôs, onde Aleph age apenas como um consultor confiável. Mas os casos interessantes estão no meio desses extremos. Talvez os processos locais do Robô A1 formem um plano de ação: Colha Flor 1. Ao mesmo tempo, chega um plano de ação originado de Aleph: Colha Flor 2. Ao invés de assumir que o plano local sempre será preferível ao plano de Aleph ou vice-versa, cada plano de ação pode ter um peso. Se Colha Flor 1 tem peso 3,4 e Colha Flor 2 tenha peso 2,6, então Colha Flor 1 é preferível, independentemente de ter sido originado em A1 ou em Aleph, e o contrário, se os pesos forem invertidos. Um modelo mais complexo pode buscar um meio termo entre planos de ações ordenados. A1 ordena Flor 1 acima de Flor 3 acima de Flor 4 acima de Flor 2. Aleph ordena Flor 2 acima de Flor 3 acima de Flor 4 acima de Flor 1. Colha Flor 3 pode, então, emergir como a melhor solução. A seleção de ações pode ser automática e não consciente ou ela pode envolver uma ambivalência consciente, tanto em A1, quanto em Aleph ou da entidade ancilar como um todo, ou em uma combinação dessas localizações. Mais uma vez, nós apenas apontamos para algumas formas pelas quais o planejamento de ações pode falhar em ser discretamente localizado. Relações mais complicadas são possíveis em agentes ancilares flexivelmente construídos.

Como a memória poderia funcionar? Para uma ação motora rápida, cada robô pode possuir uma autonomia substancial para memórias processuais básicas (e.g. como abrir frascos com amostras biológicas) e memória de curto prazo ou icônica (e.g. da posição exata das flores à medida que elas balançam ao vento). Habilidades mais complexas podem exigir interações complexas com Aleph, incluindo longos *loops* de retroalimentação e alguma diretriz *top-down*. Memórias processuais coordenadas - Robôs A1 e A2 executando um procedimento específico de colheita colaborativa, onde A1 corta as flores e A2 as ensaca -

poderiam ser distribuídas de várias formas entre A1, A2 e Aleph. Memórias semânticas e episódicas convidam à mesma diversidade de procedimentos de resolução de conflitos que a percepção, se os robôs estiverem em desacordo. Memórias “externas” podem ser armazenadas em um compartimento que ao menos um robô tem acesso rápido e confiável; ou elas podem ser escritas sobre a superfície dos corpos dos robôs, legíveis para os outros membros do pelotão; e Aleph pode não possuir quaisquer meios para saber se a representação memorial do Pelotão A foi resgatada de um compartimento, da superfície de um corpo ou do armazenamento cognitivo interno de um robô, processando essas diferentes representações memoriais da mesma forma, independente da sua fonte (compare com Clark e Chalmers, 1998).

Ao escrever e reler esse material, nós percebemos que facilmente acabamos concebendo Aleph e os robôs como mentes distintas em comunicação; e quando tentamos resistir a essa maneira de pensar, percebemos que facilmente substituímos essa ideia pela posição de que o sistema cognitivo de Aleph e dos robôs seriam subsistemas de uma mente única. O desafio que nós gostaríamos de propor a você é de conceber o sistema em nenhuma dessas duas formas ou, ao menos, não de forma *determinante* em nenhuma dessas duas formas, mas como uma arquitetura unificada o suficiente para constituir um caso que seja limítrofe entre comunicação dentro de uma mente e comunicação entre mentes.

Para facilitar a exposição, nós assumimos que, se o número de mentes for determinado, ele será um ou 201. Entretanto, outras possibilidades incluiriam haver 202 mentes (uma para cada robô, uma para Aleph e uma para o sistema ancilar como um todo), ou 10 mentes (uma para cada pelotão), ou um grande número de mentes parcialmente sobrepostas (talvez uma para cada robô, uma para cada robô mais Aleph, uma para cada pelotão e uma para cada pelotão mais Aleph).

Certamente, isso complica ainda mais a questão, criando novas oportunidades para áreas cinzentas e indeterminações.

2. A Fronteira vaga entre Introspecção e Comunicação nas Mentis Ancilares

Nós assumimos que introspecção é um processo por meio do qual alguém passa a representar sua própria mente. Necessariamente, se um processo tiver como objetivo representar a mente de outra pessoa, ele não será introspectivo (Gertler, 2000; Schwitzgebel, 2010, 2019; e, provavelmente, Kammerer e Frankish, 2023). Disso decorre que, se é indeterminado se os robôs e Aleph são mentes distintas ou subpartes de uma mesma mente, será algumas vezes correspondentemente indeterminado se as consultas de Aleph junto aos robôs constituem introspecção dentro de uma mente ou comunicação entre mentes. Em uma mente ancilar algumas atividades mentais podem ocupar uma zona cinzenta intermediária entre introspecção e comunicação.

Como isso poderia funcionar? Para simplificar, nós podemos imaginar que Aleph tem um armazenamento local de representações que ele pode acessar diretamente e que cada robô, similarmente, têm um armazenamento local de representações que eles podem acessar diretamente. Nós enfatizamos a qualificação “para simplificar”: Autoconhecimento em seres humanos ou em sistemas cognitivos sofisticados como a mente ancilar provavelmente tende a envolver elementos substanciais “indiretos” que são reconstrutivos e inferenciais (Schwitzgebel, 2012; Kammerer e Frankish, 2023). Nós vamos suspender essa questão para focar nos possíveis processos semelhantes à introspecção que poderiam ocorrer *entre* Aleph e os robôs.

Como descrito acima, na percepção, Aleph monitora os julgamentos perceptuais dos membros do pelotão (A1 representa as flores como sendo

70% azuis, A2 as representa como sendo 70% azuis, A3 as representa como sendo 50% azuis, etc.). Pode ser conveniente para Aleph continuar a monitorar os estados representacionais dos membros do pelotão A. Por exemplo, suponha que A3 receba um sinal corretivo de Aleph que deveria, se A3 estiver funcionando adequadamente, mudar a representação que A3 tem do campo para 60% azuis. Assumindo que as representações locais dos robôs contribuem de alguma forma para o seu comportamento, o comportamento de A3 será diferente do comportamento dos robôs que representam as flores como sendo 70% azuis. Mas talvez haja algum ruído no processo de correção, ou talvez Aleph não tenha recolhido todos os fatos para possibilitar uma predição, ou talvez as representações de A3 também sejam corrigidas diretamente pelos sinais dos outros membros do pelotão A, ou talvez haja uma possibilidade de mal funcionamento. O sistema Aleph pode, então, comparar sua predição sobre as representações de A3 com um sinal confirmatório de A3.

Seguindo o diagrama PID de Kammerer e Frankish (2023), o processo pelo qual Aleph rastreia as representações de A3 pode ser relativamente direto - um simples acionamento de uma representação A3-representada-60% quando A3 atualiza sua representação perceptual para 60% - ou a representação pode envolver um processo inferencial complexo. Ele pode ser bastante conceitualizado - “A3 representa 60% das flores azuis” - ou ele pode ser menos conceitual, quase analógico, como um mapa detalhado dos estados perceptuais de A3 sem rótulos interpretativos de nível superior. Ele pode ser flexível - acionado e implementado quando necessário, empregando diferentes tipos de inferências e correções dependendo da situação - ou ele pode ser inflexível e automático, surgindo sob condições pré-especificadas. Se Aleph e os robôs constituem um número indeterminado de mentes, então esses processos são indeterminados entre comunicação e introspecção. Kammerer e Frankish

podem, portanto, considerar acrescentar outra dimensão para sua descrição de variedades de introspecção: uma dimensão correspondente ao fato de o processo mirar a própria mente da pessoa (introspectivo) ou a mente de outro (não-introspectivo), com uma variedade de casos intermediários mais próximos ou distantes dos processos canonicamente introspectivos ou não-introspectivos.

Se nós assumimos que algumas configurações do sistema ancilar geram uma existência determinada e genuína tanto de uma mente sistêmica *e também* de mentes individuais para Aleph e cada um dos robôs, outra possibilidade se apresenta: introspecção no nível do sistema que é implementada pela comunicação entre mentes no nível individual dos robôs. Por exemplo, em resposta a uma pesquisa de Aleph (uma mente individual), os robôs do pelotão A (também mentes individuais) podem, cada um deles, comunicar suas representações das flores e essas repostas podem contribuir para a representação sistêmica, não apenas em Aleph, mas também na cognição da mente ancilar como um todo, na forma “75% da minha cognição do pelotão A representa o campo como 70% azul e 25% da minha cognição do pelotão A representa o campo como 60% azul.” Então, representar desacordos entre suas mentes constitutivas pode ser uma parte importante da tarefa introspectiva do sistema mais amplo. (Talvez esse desacordo seja sentido como uma ambivalência.) Portanto, uma entidade com uma mente que contenha outras mentes como partes pode introspectar seus próprios estados mentais por meios de comunicações verbais entre as mentes que ela contém.

3. A Indistinção de Algumas Mentes Ancilares

Até agora, nós assumimos que o leitor aceite que é possível criar um sistema ancilar que tenha um número indeterminado de mentes.

Nós acreditamos que essa estipulação tem alguma plausibilidade inicial se alguém aceita que interconectividade informacional (de algum tipo ou outro) é a base para a individuação de mentes e vem em graus. Conecte dois sistemas cognitivos de forma suficientemente estreita (de forma correta) e você terá uma única mente. Separe-os o suficiente, e você terá duas mentes. Plausivelmente, alguém poderia construir uma série sorites de combinações de mentes (Roelofs, 2019): Comece com um caso paradigmático de duas mentes distintas. Conecte-as juntas vagarosamente, pouco a pouco, até que no fim estejamos diante de um caso paradigmático de uma única mente. Em algum momento intermediário, parece, devem haver casos limítrofes, indeterminados entre uma ou duas mentes. Até que haja um argumento contrário, seria surpreendente se existisse uma linha clara em um ponto exato, tal que com um pouco mais de conectividade, qualquer sistema ancilar mudaria drasticamente, sem casos limítrofes, de um sistema de 201 mentes para um sistema com uma mente (ao menos se alguém aceita uma teoria da vaguidade não-epistêmica, que aceita a indeterminação).

Essa questão diz respeito não somente às mentes, mas também às pessoas, a partir do princípio de que, em geral, para sistemas cognitivamente sofisticados parecidos com o de humanos, o número de mentes é igual ao número de pessoas (entretanto, veja Schechter, 2018). Uma razão pela qual nos importamos e devemos nos importar com as fronteiras entre as mentes é que nos importamos com as fronteiras entre pessoas: Quantas pessoas morrerão caso o sistema ancilar seja destruído? Se uma promessa for feita pela boca de A3, quem está fazendo a promessa - apenas A3 ou o sistema ancilar como um todo? Se A3 se joga em um vulcão depois de instruções de Aleph, isso seria um caso de uma pessoa cometendo suicídio pelo comando de outra ou seria mais parecido com o caso de uma única pessoa que decide sacrificar um

membro? O que pensamos sobre essas questões a respeito de pessoas plausivelmente afeta como nós pensamos sobre a individuação de mentes, de forma que indeterminação sobre a individualização da mente implica indeterminação sobre a individuação de pessoas.

Se aceitamos a configuração funcionalista a partir da qual o grau de conectividade é a base da individuação de mentes e de pessoas, casos sorites parecem possíveis e a indeterminação parece ser uma consequência. Além disso, o tipo de indeterminação é curiosamente diferente de casos humanos ordinários como, por exemplo, uma pessoa morrendo lentamente, se transformando de uma mente ou uma pessoa para zero. Em mentes ancilares, o sistema pode ser determinadamente pelo menos uma pessoa, embora seja um número indeterminado de pessoas.

Nós vemos três fontes principais de resistência a essa ideia, que chamaremos de a Visão do Corpo, a Visão da Transição de Fase (veja também Schwitzgebel, 2023) e o Realismo Fenomênico Discreto. De acordo com a Visão do Corpo, nós contamos pessoas ao contar corpos. Na mente ancilar, cada robô tem um corpo. Aleph tem um corpo (provavelmente). Portanto, 201 pessoas existem, independentemente das conexões cognitivas entre elas. Note que a Visão do Corpo pode admitir pessoas indiscretas, o número de corpos pode ser indeterminado. Mas nós não discutiremos a Visão do Corpo mais profundamente. Em vez disso, vamos simplesmente assumir sua falsidade; nosso interesse é em pessoas enquanto entidades cognitivas. Meras lacunas no espaço ou na forma, que possam ser atravessadas sem demora significativa por conta da velocidade da luz, são irrelevantes para a individuação de pessoas no sentido que nos interessa (veja também Dennett, 1978; Parfit, 1984; Schwitzgebel, 2015).

A Visão da Transição de Fase observa que, algumas vezes, transições de fase repentinas interrompem o que poderia parecer um ser contínuo. A água esfria e continua esfriando, sem muitas alterações, até que, de repente, a 0,0 °C, ela se torna gelo. Uma viga vai se deformando com o aumento do peso, se deformando um pouco mais a cada grama acrescentada, até que, de repente, se quebra. Depois da transição, muita coisa é diferente. Você pode colocar uma moeda sobre o gelo. Você pode sacudir uma ponta da viga, sem sacudir a outra.

De modo similar, é possível, ao menos em princípio, que algo parecido pudesse acontecer no sistema ancilar. Integre-o lentamente, um pouco de cada vez, e, de repente, em algum ponto, com um pouquinho mais de integração, ele se altera radicalmente. Processos que operavam de uma forma, agora operam de outra maneira. Os processos que constituem percepção, memória, ação e introspecção mudam radicalmente em um piscar de olhos deixando de seguir um tipo de padrão, o padrão de 201 pessoas, para seguir um padrão muito diferente, o padrão de uma pessoa. Para contar como uma transição de fase apropriada, deve haver uma mudança substancial no funcionamento do sistema como um todo, não basta apenas ser uma mudança repentina, como a mudança de um tom de verde para um tom ligeiramente diferente não tão verde assim. Somente uma mudança repentina e substancial cria uma fronteira significativa entre as operações funcionais do caso de uma pessoa e do caso de 201 pessoas, em vez de uma nomeação diferente arbitrária ou semi-arbitrária de casos muito próximos. Com um pouquinho mais de integração, o sistema deve cair em um local muito diferente, como o pequeno toque de uma bola de basquete no aro faz com que ela caia dentro, ao invés de fora, da cesta, com todas as diferentes consequências causais que se seguem.

A Visão da Transição de Fase não é estritamente impossível. Entretanto, parece não existir razão para supor que sempre haveria tais transições de fases funcionais repentinas no curso da integração, muito menos que quaisquer transições de fase que existam devem se alinhar exatamente com as fronteiras apropriadas entre os casos de uma pessoa e de 201 pessoas, e ainda menos que todas as formas possíveis de construir o sistema ancilar teriam transições de fase bem definidas exatamente nos lugares apropriados. Além disso, se um defensor da Visão da Transição de Fase sustenta que não pode haver casos indeterminados, mesmo que por um instante, ele enfrenta o problema de que transições bem definidas que não admitem quaisquer casos limítrofes são raros entre as entidades macroscópicas da natureza. Por essas razões, na ausência de argumentos adicionais, a Visão da Transição de Fase parece ser empiricamente implausível.

Alguém poderia se esquivar do risco empírico ao aceitar a possibilidade de uma mudança metafísica bem definida sobre uma gradação suave em casos funcionalmente similares, de forma que casos muito próximos em uma série sorites que diferem somente por um único pedacinho de conectividade - sem a derivação de grandes diferenças funcionais - pode, ainda assim, ser determinadamente e claramente diferente no que diz respeito ao número de pessoas. O caso N é determinadamente bem definido como tendo 201 pessoas; o caso N+1, empiricamente quase indistinguível, determinadamente tem uma pessoa, e não existem casos intermediários.

Talvez a versão mais atraente dessa posição seja o Realismo Fenomenal Discreto, de acordo com a qual, sobre todos os processos representacionais, informacionais e cognitivos descritos nas seções 1 e 2, existem alguns fatos fenomênicos simples a respeito da pessoalidade ou subjetividade. Esses fatos fenomênicos podem ser adicionais, fatos não

físicos (como no dualismo de propriedades) ou eles podem ser fatos físicos fundados no sistema físico do ancilar. Essa proposta é plausível, ao menos em parte, devido à atratividade de tratar a fenomenalidade, a subjetividade ou o “como é ser” como necessariamente discreta, sem casos indeterminados, fronteiros (Antony, 2008; Goff, 2013; Simon, 2017; no entanto, veja Schwitzgebel, 2023 para um contra-argumento). Ou há algo como é ser o Robô A1, e o Robô A2, e o Robô A3, e assim por diante, gerando um total de 201 sistemas dos quais é correto dizer que há algo como é ser cada um deles - caso no qual haverá 201 pessoas, ou, ao menos, 201 centros subjetivos de experiências - ou não há nada como é ser esses sistemas e, ao invés disso, há algo como é ser o sistema ancilar como um todo - caso no qual há apenas uma pessoa ou centro subjetivo de experiências. Ou há uma coisa que é como é ser o todo, ou há muitas coisas que são como é ser as partes. A qualidade de “uma coisa como é ser” não pode ocorrer um número indeterminado de vezes. Fenomenalidade ou subjetividade deve ter contornos bem definidos, segundo se pensa, mesmo se os processos funcionais correspondentes tivessem uma gradação suave.

O Realismo Fenomenal Discreto também se encaixa com a proposta da unidade da consciência de Bayne (2010) e com abordagens “fenomenalistas” da introspecção descritas por Kammerer e Frankish (2023). Imagine um local de consciência no qual os estados experienciais α , β , γ , e δ ocorrem, tal que cada estado é introspeccionável pelo mesmo sujeito, Sujeito 1. Um segundo local de consciência hospeda os estados experienciais ϵ , ζ , η e θ , todos introspeccionáveis pelo Sujeito 2. Um terceiro local hospeda ι , λ , μ e ν , introspeccionáveis pelo Sujeito 3. Nenhum sujeito pode introspectar quaisquer dos estados que são introspeccionáveis por qualquer outro sujeito e não há estados experienciais flutuantes independentes de um sujeito. Relações de

“coocorrência” fenomênica são similarmente simples: Há um estado unificado no qual α , β , γ e δ ocorrem conjuntamente (da forma que suas experiências simultâneas de sons da música, gosto da cerveja e visão da banda coocorrem fenomenologicamente em um show); outro estado unificado no qual ε , ζ , η e θ coocorrem (assim como as experiências similares do seu amigo no show coocorrem em seu amigo, mas não em você); um terceiro no qual ι , λ , μ e ν coocorrem; e nenhuma relação de coocorrência de qualquer outro tipo (e.g., de α , ε e ι). Conte os sujeitos introspectivos ou locais de coocorrência fenomênica mútuas. Esse é o número de pessoas. O resultado será sempre um número inteiro.

O Realismo Fenomenal Discreto não pode estar certo se, como Kammerer e Frankish afirmam, ilusionismo sobre consciência fenomênica é verdadeiro. O Realismo Fenomenal Discreto presumivelmente requer realismo fenomênico. Mas até mesmo para não-ilusionistas, essa proposta enfrenta dois problemas.

Primeiro, ele postula, de forma deselegante, linhas ontológicas bem definidas sobre contínuos subjacentes difusos. Se o Realismo Fenomenal Discreto rejeita a Visão da Transição de Fase, então sistemas ancilares podem ser muito similares uns aos outros em todos os aspectos representacionais, informacionais e comportamentais e, ainda assim, serem radicalmente diferentes em suas relações de coocorrência introspectiva e fenomênica - um sistema abriga 201 fluxos de experiência separados, por exemplo, e o outro abriga somente um fluxo. Dada a sua similaridade representacional, informacional e comportamental, esses sistemas vão parecer virtualmente o mesmo quando vistos de fora e os robôs, Aleph e o sistema como um todo vão produzir virtualmente as mesmas saídas quando consultados sobre seus estados mentais. Não haverá o anúncio de uma experiência, similar à da transição de fase, de uma unidade repentina assim como não haverá um

“Ah! Eu me dividi em 201 partes! Me ajude, me ajude!” Adicione ou subtraia um pouquinho de conectividade e o sistema salta sobre o golfo de uma para 201 mentes e retorna novamente sem qualquer mudança psicológica significativa. Portanto, essa proposta estranhamente desassocia consciência de psicologia (compare com a crítica ao pampsiquismo em Frankish, 2021).

Segundo, o Realista Fenomenal Discreto terá dificuldades para acomodar casos de aparente sobreposição experiencial parcial. Modificando de alguma maneira a estrutura ancilar, suponha que o robô A1 e o robô A2 compartilhem muito de seus circuitos em comum. Entre eles paira uma caixa na qual a maior parte de sua cognição acontece. Talvez a caixa esteja conectada por cabos de alta-velocidade a cada um dos corpos, ou talvez, em vez disso, a informação flua através de conexões de rádio de banda larga. De qualquer modo, os processos cognitivos na caixa suspensa estão fortemente integrados cognitivamente com os corpos de A1 e A2 e os restantes de suas mentes - tão intimamente conectados quanto são em casos comuns de mentes unificadas. Embora a maior parte de sua cognição aconteça dentro da caixa, alguma cognição também ocorre no corpo individual de cada robô e não é partilhada pelo outro robô. Suponha, então, que o robô A1 tem uma experiência com um caráter qualitativo α (fundada nos processadores locais de A1), além de experiências com os caracteres qualitativos β , γ , e δ (fundadas na caixa), enquanto A2 tem experiências com caracteres qualitativos β , γ , e δ (fundadas na caixa), além de uma experiência com caráter qualitativo ϵ (fundada nos processadores locais de A2). Se a indeterminação relativa ao número de mentes é possível, talvez este não seja um sistema com um número inteiro de mentes. No entanto, o Realista Fenomenal Discreto, deve atribuir um número determinado de mentes e precisará atribuir sentido ao caso de um modo diferente.

Do nosso ponto de vista, os Realistas Fenomenais Discretos têm três opções principais: Impossibilidade, Partilha e Similaridade. De acordo com a Impossibilidade, a configuração é impossível. No entanto, não está claro por que uma configuração desse tipo deveria ser impossível, e, na ausência de um argumento mais robusto, essa opção pode ser descartada. De acordo com a Partilha, duas mentes determinadamente distintas compartilham as mesmas instâncias fenomenais das mesmas experiências com propriedades qualitativas, como β , γ e δ . De acordo com a Similaridade, duas mentes determinadamente distintas têm experiências com caracteres qualitativos β , γ e δ , mas essas experiências não são as mesmas instâncias fenomenais: as experiências β_1 , γ_1 e δ_1 de A1 são qualitativamente idênticas, mas não quantitativamente idênticas, às experiências β_2 , γ_2 e δ_2 de A2. Um desafio inicial para a Partilha é que ela parece violar a visão padrão de que as relações de coocorrência fenomênicas são transitivas (de maneira que, se α e β coocorrem fenomenicamente, e β e ε também coocorrem fenomenicamente, então o mesmo acontece a α e ε). Um desafio inicial para a Similaridade é a duplicação peculiar de instâncias experienciais. Como a caixa está conectada tanto a A1 quanto a A2, os processos que dão origem a β , γ e δ resultariam em duas instâncias de cada tipo de experiência, enquanto os mesmos processos presumivelmente dariam origem a apenas uma instância se a caixa estivesse conectada apenas a A1.

Para tornar as coisas mais desafiadoras para o Realista Fenomenal Discreto que pretende aceitar a Partilha ou à Similaridade, imagine que há um interruptor que desligará os processos em A1 e A2 responsáveis por gerar as experiências α e ε , resultando nas experiências fenomênicas totais de A1 e A2 tendo características qualitativas idênticas. Acionar o

interruptor ou colapsará A1 e A2 em uma única mente, ou não. Isso leva a um dilema tanto para a Partilha quanto para a Similaridade.

Se o defensor da Partilha sustenta que as mentes colapsam, então ele deve admitir que uma mudança relativamente pequena no campo fenomênico pode resultar em uma reconfiguração radical do número de mentes. O ponto pode ser tornado ainda mais dramático se aumentar o número de experiências na caixa e o número de robôs conectados a ela. Suponha que os 200 robôs tenham cada um 999.999 experiências originadas da caixa partilhada, e apenas uma experiência que seja qualitativamente única e localizada – talvez um círculo quase imperceptível na periferia visual esquerda de A1, um quadrado quase imperceptível na periferia visual direita de A2, etc. Se um brincalhão decidisse alternar o interruptor para frente e para trás repetidamente, na versão colapsante da Partilha, o sistema alternaria repetidamente entre ser 200 mentes e ser uma única mente, com quase nenhuma diferença na fenomenologia. Se, no entanto, o defensor da Partilha sustenta que as mentes não colapsam, ele deve admitir que múltiplas mentes distintas poderiam ter instancias fenomenais idênticas e fundadas nos mesmos processadores cognitivos. Essa visão levanta a questão sobre a base ontológica da individuação das mentes; em algumas concepções de sujeito, essa visão pode não ser sequer coerente. Parece postular sujeitos com diferenças metafísicas, mas não fenomenológicas, o que contraria o espírito geral do realismo fenomenal sobre as mentes.

O defensor da Similaridade enfrenta problemas análogos. Se ele sustenta que o número de mentes colapsa para um, então, tal como o defensor da Partilha, ele deve permitir que uma mudança relativamente pequena no campo fenomênico possa resultar em uma mudança radical no número de mentes. Além disso, deve permitir que experiências distintas, meramente idênticas em termos de tipo, se tornem, de

alguma forma, uma mesma experiência quando se liga um interruptor que praticamente não altera a fenomenologia do sistema. Mas se sustenta não haver colapso, então depara-se com a possibilidade incomoda de existirem múltiplas mentes distintas com experiências qualitativamente idênticas, mas numericamente distintas, resultantes dos mesmos processadores cognitivos. Isto parece ser uma inflação fenomenal ontologicamente não-parcimoniosa.

Para recapitular: Se aceitarmos que a individuação de mentes e pessoas está fundada sobre fatos acerca da interconexão funcional que podem, em princípio, ser organizados numa série de sorites com indeterminação no meio, então somos forçados a uma de duas opções: Ou aceitamos que o número de mentes e pessoas pode ser indeterminado, ou insistimos numa linha metafísica clara apesar da hipotética possibilidade de construção de uma série de sorites. A segunda opção divide-se em duas sub-opções: Se houver uma linha metafísica clara, ela pode ou não envolver uma grande diferença funcional. Se envolver uma grande diferença funcional (a Visão da Transição de Fase), então as configurações intermediárias na série de sorites devem ser funcionalmente instáveis, colapsando rapidamente em configurações mente/pessoa funcionalmente distintas e discretamente contáveis. Parece haver poucas razões para pensar, empiricamente, que isto tem de ser assim. Se não envolver uma grande diferença funcional, então parece que a melhor opção é o Realismo Fenomenal Discreto. Mas o Realismo Fenomenal Discreto parece exigir linhas metafísicas estranhamente claras na ausência de grandes diferenças funcionais. Além disso, diferentes variações do Realismo Fenomenal Discreto parecem enfrentar sérios problemas para explicar a possibilidade de similaridade qualitativa entre mentes.

4. De volta à Terra: Países, Personalidade Múltipla, Casos de Cérebro Dividido e Gêmeos Craniópagos

Se estamos dispostos a abandonar a contabilidade discreta e aceitar a possibilidade de processos mentais semelhantes à introspecção para casos não convencionais da ficção científica, então talvez tenhamos que fazê-lo também para alguns casos na Terra.

Em outro trabalho (Schwitzgebel 2015), um de nós argumentou que não é absurdo pensar que os Estados Unidos, concebido como uma entidade concreta com pessoas como partes, literalmente possui um fluxo de consciência além das experiências dos cidadãos e residentes que o constituem. Essa entidade representa, metarrepresenta, envolve-se em processamento de informações massivas e complexas e responde de maneira inteligente ao seu ambiente. Uma preocupação em relação a esse ponto de vista é que, se aceitarmos que os Estados Unidos estão conscientes, também teremos que aceitar que Luxemburgo e a *Microsoft Corporation* estão conscientes; e se aceitarmos que Luxemburgo e a *Microsoft Corporation* são conscientes, também teremos que aceitar que Delaware e a Universidade da Califórnia, Riverside são conscientes; e assim por diante, até uma posição ainda mais aparentemente absurda do que a ideia de que os Estados Unidos está consciente. Se a consciência não é uma questão de tudo ou nada e as entidades conscientes nem sempre precisam ser discretas e contáveis, então talvez a consciência desapareça gradualmente, em vez de repentinamente, à medida que as entidades de grupo se tornam menos parecidas com uma mente única e interconectada. Se as mentes conscientes podem estar presentes, mas indiscretas, em sistemas auxiliares hipotéticos, então talvez elas possam estar presentes, mas indiscretas, em grupos realmente existentes na Terra.

Alguns casos humanos também podem envolver indeterminação em relação ao número de mentes e pessoas, bem como na presença de introspecção versus comunicação. Casos problemáticos bem conhecidos incluem Transtorno Dissociativo de Identidade ou Personalidade Múltipla (Braude 1995; Radden 1996; Maiese 2016), casos de cérebro dividido (Schechter 2018), gêmeos craniópagos com regiões cerebrais sobrepostas (Cochrane 2021) e talvez alguns animais não humanos, como o polvo, com processamento cognitivo substancialmente desunificado (Godfrey-Smith 2016). Diante dos gêmeos Hogan, ou de Jekyll-Hyde, a resposta correta para o número de mentes ou pessoas pode não ser nem um nem dois, mas algo indeterminado entre eles. Os gêmeos Hogan e Jekyll-Hyde podem às vezes se envolver em cognição que não é nem determinadamente introspecção dentro de uma mente nem determinadamente comunicação entre duas mentes. Se aceitarmos isto, então talvez até as mentes típicas sejam menos discretamente individuadas do que tendemos a pensar. Talvez sejamos todos um pouco imprecisamente delimitados, desunidos e plurais.⁴

Referências

- ANTONY, Michael. Vagueness and the metaphysics of consciousness. *Philosophical Studies*, 128, 515-538, 2006.
- BAYNE, Tim. *The unity of consciousness*. Oxford University Press, 2010
- BRAUDE, Stephen. *First person plural*. Rowman & Littlefield, 1995.
- CLARK, Andy; CHALMERS, David. The extended mind. *Analysis*, 58, 7-19, 1998.

⁴ Agradecemos a Rotem Herrmann pelas considerações sobre memória em sistemas ancilares (veja também Schwitzgebel e Herrmann, 2016) e a Nichi Yes pelas reflexões sobre identidade pessoal em Transtorno Dissociativo de Identidade. Por comentários úteis sobre o manuscrito, agradecemos a Brice Bantegnie, Tim Bayne, Rotem Herrmann, François Kammerer e Luke Roelofs.

COCHRANE, Tom. A case of shared consciousness. *Synthese*, 199, 1019-1037, 2021.

DENNETT, Daniel C. Where am I? In D.C. DENNETT, *Brainstorms*, MIT Press, 1978.

FRANKISH, Keith. Panpsychism and the depsychologization of consciousness. *Aristotelian Society Supplementary Volume*, 95, 51-70, 2021.

GERTLER, Brie. The mechanics of self-knowledge. *Philosophical Topics*, 28 (2), 125-146, 2000.

GODFREY-SMITH, Peter. *Other minds*. Farrar, Straus and Giroux, 2016.

GOFF, Philip. Orthodox property dualism + the Linguistic Theory of Vagueness = Panpsychism. In R. BROWN, ed., *Consciousness inside and out*. Springer, 2013.

KAMMERER, François; FRANKISH, Keith. What Forms Could Introspective Systems Take? A Research Programme. *Journal of Consciousness Studies*, v. 30, n. 9-10, p. 13-48, 2023.

LECKIE, Ann. *Justiça Ancilar*. São Paulo: Aleph, 2018.

MAIESE, Michelle. *Embodied selves and divided minds*. Oxford University Press, 2016.

PARFIT, Derek. *Reasons and persons*. Oxford University Press, 1984.

RADDEN, Jennifer. *Divided minds and successive selves*. MIT Press, 1996.

ROELOFS, Luke. *Combining minds*. Oxford University Press, 2019.

SCHECHTER, Elizabeth. *Self-consciousness and "split" brains*. Oxford University Press, 2018.

SCHWITZGEBEL, Eric. Introspection. *Stanford Encyclopedia of Philosophy*, 2019.

SCHWITZGEBEL, Eric. Introspection, what? In D. SMITHIES and D. STOLJAR, eds., *Introspection and consciousness*. Oxford University Press, 2012

SCHWITZGEBEL, Eric. If materialism is true, the United States is probably consciousness. *Philosophical Studies*, 172, 1697-1721, 2015.

SCHWITZGEBEL, Eric. *Borderline consciousness, when it's neither determinately true nor determinately false that experience is present*, 2023. Disponível em: <http://faculty.ucr.edu/~eschwitz/SchwitzAbs/BorderlineConsciousness.htm>.

SCHWITZGEBEL, Eric; HERRMANN, Rotem. Possible architectures of group minds: Memory. Blog post at *The Splintered Mind* (Jun 14), 2016. <https://schwitzsplinters.blogspot.com/2016/06/possible-architectures-of-group-minds.html>.

SIMON, Jonathan. Vagueness and zombies: Why 'phenomenally conscious' has no borderline cases. *Philosophical Studies*, 174, 2105–2123, 2017.

RELAÇÃO DE AUTORES

Anderson Luiz do Vale Fonseca - Doutorando na Filosofia pela Universidade Federal do Ceará. Mestre em Filosofia pela Universidade Federal de São João del-Rei.

Aryel C. Goes – Doutorando em Evolução, Ecologia e Biologia Organismal na Universidade Estadual de Ohio. Mestre em Biologia Celular, Molecular e Microbiologia pela Universidade Estadual Paulista – Rio Claro.

Daniel Friedman – Membro da Diretoria do Instituto de Inferência Ativa. Doutor em Biologia pela Universidade Stanford.

Eric Schwitzgebel – Professor de Filosofia na Universidade da Califórnia – Riverside. Doutor em Filosofia pela Universidade da Califórnia – Berkeley.

José Carlos Camillo - Doutorando em filosofia na Universidade Federal de Goiás e Doutorando Visitante na Ruhr-Universität Bochum. Bolsista da FAPEG e do DAAD. Mestre em Filosofia pela Universidade Federal de Mato Grosso.

Luiz Paulo Da Cas Cichoski - Professor de Filosofia na Universidade Federal de Mato Grosso. Doutor em Filosofia pela Pontifícia Universidade Católica do Rio Grande do Sul.

Luke Roelofs – Professor de Filosofia na Universidade do Texas – Arlington. Doutor em Filosofia pela Universidade de Toronto.

Maria Luiza Iennaco – Doutoranda em Filosofia na Universidade de São Paulo. Mestre em Filosofia pela Universidade de São Paulo.

Robert Rupert – Professor de Filosofia na Universidade do Colorado – Boulder. Doutor em Filosofia pela Universidade de Illinois – Chicago.

Sophie Nelson - Doutoranda em Filosofia na Universidade de Nova York. Bacharel em Filosofia e Psicologia pela Universidade Oberlin.

Valdenor Monteiro Brito Júnior - Professor Substituto de Filosofia na Universidade Federal do Paraná. Doutor em Filosofia pela Universidade Federal de Santa Catarina.



A Editora Fi é especializada na editoração, publicação e divulgação de produção e pesquisa científica/acadêmica das ciências humanas, distribuída exclusivamente sob acesso aberto, com parceria das mais diversas instituições de ensino superior no Brasil e exterior, assim como monografias, dissertações, teses, tal como coletâneas de grupos de pesquisa e anais de eventos.

Conheça nosso catálogo e siga as nossas páginas nas principais redes sociais para acompanhar novos lançamentos e eventos.



www.editorafi.org
contato@editorafi.org

É POSSÍVEL QUE COLETIVOS TENHAM UMA MENTE? MAIS ESPECIFICAMENTE, A HIPÓTESE DE QUE UMA COLÔNIA DE FORMIGAS SEJA UMA ENTIDADE CONSCIENTE É RAZOÁVEL? SE A RESPOSTA FOR POSITIVA, QUAIS SERIAM AS IMPLICAÇÕES FILOSÓFICAS? AO OLHARMOS PARA O MINÚSCULO UNIVERSO DAS FORMIGAS NOS MARAVILHAMOS COM A POSSIBILIDADE DE UMA MENTE EMERGIR DE SUA EMARANHADA TEIA DE CONEXÕES E QUE ESTRUTURAS SIMILARES POSSAM TAMBÉM SER SISTEMAS COGNITIVOS. ESTE LIVRO FOI ORGANIZADO COM O OBJETIVO DE REUNIR DIFERENTES PONTOS DE VISTA SOBRE ESSE TEMA, APRESENTANDO UMA VISÃO AMPLA E ROBUSTA DAS INVESTIGAÇÕES SOBRE CONSCIÊNCIA COLETIVA. GRANDE PARTE DAS CONTRIBUIÇÕES SEGUIRÁ A QUESTÃO CENTRAL DO LIVRO, APRESENTANDO UMA ABORDAGEM NATURALISTA, BIOLÓGICAMENTE CENTRADA, SOBRE CONSCIÊNCIA COLETIVA. NO ENTANTO, ALGUNS DOS TEXTOS TRAZEM PERSPECTIVAS MAIS TRADICIONAIS, FOCADOS MAIS NAS EXPERIÊNCIAS HUMANAS, TENDO, PORTANTO, UMA ABORDAGEM MAIS ANTROPOLÓGICA DO QUE ZOOLOGICA. O LEITOR TAMBÉM ENCONTRARÁ UMA DIVERSIDADE DE POSICIONAMENTOS SOBRE NOSSA QUESTÃO CENTRAL. AINDA QUE A MAIORIA DOS AUTORES ESBOCE SIMPATIA PELA TESE DA CONSCIÊNCIA COLETIVA, BUSCAMOS INCLUIR VOZES MAIS HESITANTES E ATÉ MESMO CONTRÁRIAS A ELA.

