

# Processos Cognitivos e Mente Estendida: uma metáfora neofuncionalista?

## *Cognitive Processes and Extended-Mind: a neofunctionalist metaphor?*

Léo Peruzzo Junior<sup>1</sup>  
Amanda Luiza Stroparo<sup>2</sup>

**Resumo:** Este artigo pretende analisar como a hipótese da mente estendida, por um lado, redesenha o conceito de processos cognitivos a partir de uma crítica ao paralelismo físico/mental e, por outro, se ampara em argumentos funcionalistas para considerar outros níveis do organismo e do ambiente. Para tanto, inicialmente, mostramos como a dicotomia físico/mental constrói uma imagem epistemologicamente confusa, especialmente quando se considera o argumento da causação mental. A tese de que os processos mentais interferem nos fenômenos físicos, ou vice-versa, é um legado da tradição cartesiana. Posteriormente, argumentamos que o conceito de mente estendida não explica adequadamente o uso dos artefatos, inferindo um prognóstico evolutivo na espécie humana. Por fim, identificamos alguns dilemas – epistemológicos e morais – entre os processos cognitivos e propostas da Inteligências Artificial.

**Palavras-chave:** Mente Estendida; Cognição; Funcionalismo; Inteligência Artificial.

**Abstract:** *This article intends to analyze how the extended mind hypothesis, on one hand, redraws the concept of cognitive processes based on a critique of physical/mental parallelism and, on the other hand, is grounded on functionalist arguments to consider other levels of the organism and the environment. With this purpose, initially, we indicate how the physical/mental dichotomy builds an epistemologically confused image, especially when considering the mental causation argument. The thesis that mental processes interfere with physical phenomena, or vice versa, is a legacy of the Cartesian tradition. Later, we argue that the concept of an extended mind does not adequately explain the use of artifacts, inferring an evolutionary prognosis in the human species. Finally, we identified some dilemmas – epistemological and moral – between the cognitive processes and proposals of Artificial Intelligence.*

---

<sup>1</sup> Professor do Programa de Pós-Graduação em Filosofia Mestrado e Doutorado da Pontifícia Universidade Católica do Paraná / PUCPR, do Centro Universitário Franciscano do Paraná - UNIFAE/PR e da Faculdade Vicentina - FAVI. Doutor em Filosofia pela Universidade Federal de Santa Catarina - UFSC. Editor-Chefe da Aurora Journal of Philosophy (Qualis A2). Além do Brasil, realizou conferências na Espanha, Portugal, Itália, Coréia do Sul, México, Peru e Argentina. Pesquisa os seguintes temas: Epistemologia e Representação Mental do Conteúdo Moral; Linguagem e Intencionalidade; Transumanismo e Cognição; Filosofia do Direito. E-mail: leoperuzzo@hotmail.com

<sup>1</sup> Graduada em Psicologia pela Pontifícia Universidade Católica do Paraná em fevereiro de 2018. Mestranda em Filosofia pela mesma instituição, em Filosofia da Mente e Pós-Humanismo. Possui interesses nos seguintes campos do conhecimento: Filosofia da Mente, Ciências Cognitivas, Pós-Humanismo, Psicologia Analítica, Logoterapia e Educação. Contato: als.stroparo@hotmail.com

<sup>2</sup> Graduada em Psicologia pela Pontifícia Universidade Católica do Paraná em fevereiro de 2018. Mestranda em Filosofia pela mesma instituição, em Filosofia da Mente e Pós-Humanismo. Possui interesses nos seguintes campos do conhecimento: Filosofia da Mente, Ciências Cognitivas, Pós-Humanismo, Psicologia Analítica, Logoterapia e Educação. Contato: als.stroparo@hotmail.com

**Keywords:** *Extended Mind; Cognition; Functionalism; Artificial Intelligence.*

## **1. Do dualismo físico-mental à uma teoria dos Processos Cognitivos**

Tradicionalmente, os processos cognitivos foram interpretados a partir do paralelismo físico/mental. A dicotomia cartesiana serviu para alimentar problemas epistemológicos dos mais diversos, entre os quais o problema da “causação mental” e da analogia entre seres humanos e máquinas, bem descrita por Wilks (1975) ao criticar Putnam e Clarke. A hipótese de que os processos mentais interferem nos fenômenos físicos, ou vice-versa, é um legado da tradição cartesiana, em especial que a subjetividade seria uma qualidade não redutível aos demais eventos que ocorrem e possuem extensionalidade. Nesse sentido, Barack (2019, p.63) indica que nas ciências cognitivas há, particularmente, duas abordagens para a análise da cognição: o dinamismo e o computacionalismo. Na primeira, a cognição é analisada nos termos de subsistemas em interação do cérebro, corpo e ambiente; na segunda, compreende-se os fenômenos cognitivos em termos de tais partes e suas funções. De qualquer forma, modelos para a compreensão da cognição poderiam auxiliar, por exemplo, no modo como intervimos no ambiente a fim de traçar prognósticos diante da época que, não raros pesquisadores, têm chamado de Antropoceno (Santana, 2019).

A acessibilidade ao conteúdo cognitivo, tomado como interno ou privado, foi considerada uma espécie de artefato cujo monopólio não passava da figura de seu possuidor. Em referência à hipótese anterior, Ryle realiza um empreendimento filosófico a fim de rebater o homúnculo cartesiano, ao qual se referiu também como “teatro cartesiano”. O argumento cartesiano não era apenas falso, como dotado de uma peculiaridade por não considerar a expressão linguística dos termos que utilizamos para a expressão do mental. Segundo Ryle (2009, p.1), a doutrina oficial, pode ser descrita assim: “Com duvidosas exceções de idiotas e bebês, todo ser humano tem um corpo e uma mente. Alguns preferem dizer que todo ser humano é um corpo e uma mente. Seu corpo e sua mente são normalmente reunidos, mas após a morte do corpo, sua mente pode continuar existindo e funcionando”.

Desse modo, a percepção interna seria uma espécie de método para a avaliação contínua dos dados que aí estariam armazenados. Com isso, Ryle conclui que “termos” e “ações” seriam absolutamente distintos: termos possuem uma equivalência conceitual

referencial ao conteúdo mental; ações são um movimento abarcado também pelos outros comportamentos que fazem parte da expressão. Ryle (2009, p. 2-4) quer expurgar, em linhas gerais, como ele mesmo observa, algumas teses que parecem gerar continuamente problemas metafísicos e epistemológicos, entre as quais se a mente está no espaço, se suas operações estão sujeitas às leis da mecânica e se o seu funcionamento é testemunhado por outros observadores (Peruzzo Júnior, 2019).

O paradigma cartesiano é uma espécie de retrato que perpassa continuamente a história da filosofia da mente e das ciências cognitivas (Pirenne, 1950; Watson, 1951; Bloor, 1970). Dentre as questões que acabamos de mencionar, há outra que também merece particular atenção: a mente ocupa um lugar privilegiado no mundo? Hofweber (2019, p. 771), no ensaio *Idealism and the Harmony of Thought and Reality*, afirma que o *idealismo* se tornou uma “[...] grande visão metafísica do lugar das mentes ou do mental na realidade. [...] O idealismo é uma terceira resposta [a primeira, seria uma resposta naturalista e, a segunda, teísta] para a questão de como mentes e matéria se relacionam entre si”. Assim, embora o problema mostrado pelo idealismo pareça ser significativo, Hofweber (2019, p. 771-772) sustenta que apenas um “*strong idealism*” poderia nos dar “[...] uma terceira resposta à grande questão sobre nosso lugar no mundo [...]” e porque “[...] talvez sejamos as únicas criaturas que pensam a realidade e que se entendem partes dela”.

De qualquer modo, a hipótese apresentada por Hofweber acerca do idealismo, parcialmente abandonada, nos parece central quando analisamos o problema da cognição. Supor que tenhamos algum lugar ou não enseja apenas uma derrocada metafísica. Por isso, a hipótese da mente estendida parece dissolver determinados obstáculos filosóficos ao defender que os estados mentais dependem fortemente da estrutura corporal associada ao funcionamento dos diversos níveis do organismo (Clark, 1997; Clark & Chalmers, 1998; 2008). Ao mesmo tempo, a teoria da mente estendida admite a presença de artefatos na constituição dos processos mentais, tornando requerente que partes desses processos ocorram em outro substrato. Embora a imagem possa parecer simplificadora, essa posição acaba mostrando que algumas propriedades dos sistemas e processos cognitivos não podem ser explicadas apenas por uma descrição neural, tornando-se uma posição oposta ao dualismo e ao fisicalismo. Nosso argumento, neste trabalho, é mostrar os horizontes conceituais da cognição estendida para, então, sustentar que ela é um novo tipo de funcionalismo.

Delimitar e compreender o funcionamento da cognição, portanto, pode não revelar apenas os componentes do modo como realizamos o acesso ao mundo, mas também associar-se aos projetos que permitem replicar a cognição humana para o desenvolvimento de processos e tecnologias artificiais. Isso significa, preliminarmente, por um lado, aceitar que a cognição não é simplesmente uma atividade cerebral isolada dos demais níveis do organismo e do ambiente e, por outro, que sua arquitetura dinâmica dificilmente poderá ser axiomatizada por um modelo padrão. Contudo, é possível um gerenciamento de processos artificiais, supondo que estes gozem de similaridade com a cognição humana, se esta última ainda possui desvantagens metodológicas no modelo de sua compreensão?

Edmond Awad e outros pesquisadores publicaram um estudo intitulado *The Moral Machine Experiment*. A partir de uma plataforma que reuniu 40 milhões de decisões, em 10 idiomas e 233 países e territórios, o mesmo procurou avaliar as preferências morais globais, além de variações éticas transculturais, princípios globais e socialmente aceitáveis para a ética das máquinas. Segundo Edmond Awad et al. (2018, p.59),

[...] precisamos avaliar as expectativas sociais sobre como os veículos autônomos devem resolver dilemas morais. [...] pode-se testar se as pessoas preferem poupar muitas vidas ao invés de poucas; ou se as pessoas preferem poupar pedestres que atravessam legalmente do que os pedestres fora do passadiço; ou ainda alguma outra preferência ou uma simples combinação de duas ou três dessas preferências.

Em que pese as hipóteses do projeto, há outra ainda mais complexa descrita pelos autores como “*universal machine ethics*”:

[...] o segundo desafio torna os requisitos do tamanho da amostra ainda mais assustador: se quisermos progredir rumo à ética universal das máquinas [...] precisamos de uma compreensão detalhada de como diferentes indivíduos e países podem diferir em suas preferências éticas (Awad et al., 2018, p.59).

Outra referência analisada pelos pesquisadores parte de premissas apresentadas e elaborados pelo *Bundesministerium für Verkehr und digitale Infrastruktur* (Ministério Federal de Transporte e Infraestrutura Digital) da Alemanha, mais especificamente no documento intitulado *Ethics Commission Automated and Connected Driving*. Na regra 7, o documento afirma que em situações perigosas que se revelam inevitáveis, apesar de todas as preocupações tecnológicas tomadas, “[...] a proteção da vida humana goza de prioridade máxima no equilíbrio entre os interesses protegidos. Assim, dentro das restrições do que é tecnologicamente viável, os sistemas devem ser programados para

aceitar danos a animais ou propriedades [...]”. A cognição, nesses termos, parece estar diluída na pressuposição de que dilemas morais possam servir-se de um parâmetro para explicar seu funcionamento, permitindo inclusive a instanciação de níveis de decisão em seres humanos ou máquinas em momentos de conflito. Contudo, a conclusão otimista de Edmond Awad e colaboradores (2018, p. 63) – “[...] nunca na história da humanidade permitimos que uma máquina autônoma decida quem deve viver e quem deve morrer, numa fração de segundos, sem supervisão em tempo real” – encena um horizonte onde os processos cognitivos são encarados como um conjunto de variáveis que podem ser cruzadas e mensuradas a partir de uma proposta de “ética universal”. Mas, efetivamente, como os processos cognitivos se “encaixam”, “sentem” e “percebem” o mundo?

## **2. A naturalização da Cognição e o Mito da Percepção Objetiva**

Ao assumirmos a hipótese preliminar de que a cognição e os processos cognitivos são atrelados à dinâmica fisiológica cerebral, devemos também aceitar a existência de um dilema quando consideramos a percepção e/ou recepção dos fenômenos exteriores, geralmente desconsiderados. Alguns autores (Nagel, 1974; Searle, 1980) têm sustentando que os eventos exteriores são irrelevantes ou insuficientes para a avaliação da qualidade do conteúdo mental; outros (Aranyosi, 2011; Churchland, 2013; Hochstein, 2017), ao contrário, têm argumentado que o conflito epistêmico entre mental e físico é de natureza aparente, uma vez que há uma expectativa gramatical escondendo os reais processos fisiológicos e químicos que acontecem no cérebro.

De qualquer maneira, tanto o primeiro grupo de autores quanto o segundo não utilizam de um dado importante da neuropsicologia cognitiva, a saber, a interação entre a investigação de processos cognitivos normais e danificados para inferir suas conclusões. Segundo Stone e Davies (1993, p. 592), “os déficits aparentes em pacientes com distúrbios cognitivos adquiridos é mais proveitosa se for baseada no desenvolvimento de modelos de processo normal” e na “evidência convergente de estudos de indivíduos normais e com lesão cerebral”. Alguns pacientes, por exemplo, após lesões cerebrais, não são capazes de reconhecer faces em fotografias, não conseguem estabelecer o significado de palavras, apresentam alteração de comportamento ou não conseguem associar sons às palavras (Stone & Davies, 1993, p. 593-596).

Os episódios descritos por Stone e David (1993), entre outros autores, convergem especialmente para duas perspectivas nas ciências cognitivas e, conseqüentemente, em filosofia da mente. A primeira destas é que os processos cognitivos são eventos naturais que não podem ser tomados em paralelo aos cerebrais. Assim, o argumento da naturalização da cognição sustenta-se pelas premissas de que esta possa sofrer danos quando afetada, uma vez que seu conteúdo depende também da interação com o ambiente circundante. A segunda perspectiva é mais simples, porém mais severa. Trata-se da dissociação de que o conteúdo mental não possa ser reduzido ao corolário de informações que são processadas por este sistema.

Nesse sentido, as ciências cognitivas têm mostrado que a formação de conceitos fundados na percepção é a base de conceitos que fornecem explicações causais de entidades ou propriedades (Krellenstein, 1995). Os modelos teóricos que seres humanos usam para descrever o mundo, portanto, são moldados por uma extensão analógica daquilo que é observado. Algumas pesquisas sobre o desenvolvimento cognitivo, como apontado por Michael Vlerick (2019, p. 7-8), mostram que a “[...] noção de ‘objetividade’ está geneticamente conectada e presente no nascimento. [...] Perceber e imaginar os objetos é, portanto, um produto contingente do cérebro e dos sentidos que a evolução nos dotou”, sendo que “[...] a base perceptiva das entidades modeladas é produto de nosso aparato sensorial – mais precisamente, nosso aparato visual”.

É naturalmente evidente que a interação da cognição com o ambiente também precisa explicar não apenas a potencialidade da capacidade, mas também por que alguns elementos da realidade, embora observados e percebidos, não possuem qualquer importância para a imagem e consciência que temos do mundo. A ideia de uma “percepção objetiva” do mundo é uma doutrina enganada. Quando as pessoas conscientemente movem os olhos, distorcendo movimentos oculares sacádicos, sua imagem mental é interrompida (Vlerick, 2019), sugerindo, portanto, que a imaginação mental está intimamente conectada à visão, diferentemente de outras espécies de animais que utilizam mais eficazmente o senso olfativo ou auditivo.

Em outras palavras, não há um mundo em si, como poderia objetar um realista metafísico, do mesmo modo como não há uma percepção, imagem ou representação mental *ipso facto* à natureza, como argumentaria um mentalista ou neocartesiano. Um conceito físico que utilizamos fornece um substrato diferente para modelar o mundo, independente dos modelos experimentais, lógicos, matemáticos ou teóricos que utilizamos (Barbour, 1974, pp. 29-30 *apud* Vlerick, 2019, p. 9) para observar um fenômeno. Por essa razão, a hipótese da

mente estendida, de algum modo, parece revitalizar o projeto funcionalista, sendo uma espécie de alternativa ao materialismo reducionista e ao dualismo. Segundo o funcionalismo, “os estados mentais são definidos e caracterizados pelo papel funcional que desempenham no caminho entre a recepção dos dados e a resposta do organismo de um organismo ou sistema [...]” (Teixeira, 2008, p. 124). Mas, como podemos caracterizar os artefatos utilizados pela percepção para “ver” e “conhecer” o mundo?

### **3. A hipótese da mente estendida**

A proposta da mente estendida pode se tornar uma alternativa, visto que, conforme mencionado anteriormente, pretende ampliar o entendimento da cognição para além das estruturas orgânicas do cérebro, embora ainda concebendo seu papel crucial na relação com operações realizadas no ambiente e entre outros pares. Andy Clark (2004), em *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*, defende que ser “ciborgue” é uma característica intrínseca do ser humano, presente em toda a história e responsável por grande parte da evolução da espécie, especialmente no que tange à utilização de ferramentas e tecnologias para a sobrevivência, a otimização de operações e a resolução de problemas em geral.

O termo “externalismo ativo”, cunhado no artigo *The Extended Mind*, por Clark e David Chalmers (1998), também ilustra essa perspectiva, na medida em que o ambiente “externo” ao aparato biológico do ser humano exerce função ativa nos processos cognitivos. Mente e ambiente se tornam um “[...] sistema acoplado, que pode ser entendido como um sistema cognitivo legítimo” (Clark & Chalmers, 1998, p. 8), visto que mudanças no ambiente poderiam afetar profundamente os processos cognitivos e, por conseguinte, os comportamentos. A relevância causal de processos cerebrais é a mesma dos aparatos externos.

Destarte, posto que “[...] é de nossa natureza humana básica anexar, explorar e incorporar profundamente materiais não biológicos em nossos perfis mentais” (Clark, 2004, p. 198), não apenas o cérebro, o corpo e o ambiente ao redor são partes relevantes do sistema cognitivo, mas também os artefatos que o ser humano vem a utilizar como ferramentas auxiliares, suportes de sua atuação no mundo (Clark, 2001). Uma das perguntas essenciais acerca dos processos cognitivos não seria, então, concluir se há um

amálgama entre algo humano e algo não-humano, mas sim investigar como ocorre tal relação.

De acordo com Clark (1990), para esclarecer mecanismos através dos quais ocorrem tais relações, os estudos de inteligência artificial podem ser de grande valia, não porque poderiam necessariamente replicar a cognição humana de maneira exata e fiel, e sim por se constituírem como modelos explicativos. Dessa forma, embasando-se em uma perspectiva evolutiva – a qual encontra grande inspiração nos escritos de Daniel Dennett (1988), como a obra *The evolution of consciousness* –, Clark (1990) propõe que estudos em inteligência artificial que partirem de processos cognitivos básicos, como a percepção e a aquisição de linguagem, ao contrário de processos de “alto nível”, como aqueles de raciocínio lógico, poderão lançar luz ao funcionamento de processos cognitivos e aos modos através dos quais estes acoplam elementos do meio externo.

O eventual erro de projetos em inteligência artificial está em supor que apenas um modelo é capaz de simular a mente humana, como é o caso de estudos clássicos com computadores que têm como tarefa jogar xadrez ou resolver equações matemáticas. Tais processos lógicos também são importantes para o sucesso da cognição, porém a implementação de tais mecanismos de maneira isolada é incapaz de simular processos cognitivos que certamente teriam sido alcançados através de um processo evolutivo, tais como acoplar com o ambiente para solucionar problemas com o mínimo de esforço isolado possível, a flexibilidade e a capacidade de improvisar, o entendimento de analogias e metáforas, a interpretação de elementos linguísticos de acordo com o contexto, dentre outros. Nesse ponto, Clark (1990) explicitamente concorda com John Searle em seu experimento mental do Quarto Chinês, embora discorde de que o substrato biológico cerebral seja o responsável pela particularidade da cognição humana em sua performance.

Mentes, portanto, “[...] não são dispositivos de raciocínio lógico desprovido de corpo [...]”, seriam muito menos um “centro de controle” responsável por organizar e manipular informações e comportamentos. Ao contrário, como se pode pressupor de uma perspectiva evolucionista, a cognição seria, em verdade, uma “cognição corporificada” direcionada de acordo com “[...] demandas de sistemas produtores de comportamento em tempo real [...]”, sempre atuando em conjunto com o ambiente em seu próprio favor (Clark, 1997, p. 23).



Completar padrões (*pattern completion*) – juntamente com a característica de se conectar ao meio externo e de ser um sistema de “auto-organização” – seria uma das habilidades mais essenciais da cognição, que tem como uma de suas características mais fundamentais a predição de ações, comportamentos, informações etc. (Clark, 1990; 2016), uma vez que o processamento preditivo oferece o parceiro neurocomputacional perfeito para trabalhos recentes acerca da mente corporificada. Nessa visão, o sistema cognitivo é compreendido a partir de uma perspectiva do cérebro como motor associativo, sendo as interações ambientais uma série repetida de cálculos simples de completar padrões.

Clark (2016) se preocupa, no entanto, em definir o que tornaria o processamento cognitivo humano supostamente tão particular quando comparado a outros mamíferos, uma vez que a característica de explorar o ambiente para otimizar processos pode ser encontrada em outras espécies. Para o autor, uma das possíveis respostas é a utilização, não somente do ambiente, mas de práticas e aprendizagens decorrentes de interações sociais. A linguagem, a cultura e as interações sociais surgem nesse cenário como possíveis artefatos de suporte (*scaffolding*).

Assim, tanto o “teatro cartesiano” quanto a hipótese fisicalista são perspectivas às quais Clark se opõe. Mesmo quando falamos sobre artefatos, não estaríamos distinguindo entre um ser dotado de habilidades, capaz de incorporar elementos do meio externo, o que resultaria em uma proposta dualista ou mentalista da mesma maneira. Clark (2004, p. 192) esclarece que

[...] não existe um usuário constituído por informações em relação a quem todo o resto são somente ferramentas. É tudo ferramenta até o fim. Somos apenas as aglomerações complexas e mutáveis de “nossas próprias” ferramentas internas e externas para o pensamento. Nós somos nossos melhores artefatos, e sempre fomos.

Artefatos parecem ser definidos, por conseguinte, como ferramentas e processos inicialmente externos ao aparato biológico que depois foram “incorporados” ou “integrados” à cognição humana por oferecer uma operação que desempenha uma função necessária à vida e sobrevivência diárias. Clark e Chalmers (1998) comparam a função do *notebook* para resgatar informações com a função do cérebro para armazenar dados, ambos contribuindo com a memória. Nessa via, parece plausível admitir que se trata de uma teoria funcionalista; contudo, não mais de um caráter somente informacional e

abstrato, posto que propõe a ampliação da cognição e, em consequência, a impossibilidade de interpretar a mente como um mero intercâmbio de dados dotados de funções e cujo funcionamento seria redutível ao cérebro.

Menary (2007) é um dos autores que realiza uma leitura funcionalista da mente estendida. Como nos parece, porém, a hipótese em questão enfrenta os mesmos problemas epistemológicos daqueles decorrentes do próprio funcionalismo: a causação mental ou superveniência mente/corpo, o uso dos termos “conceitos” e “propriedades mentais” e o modo como propriedades não-mentais realizam propriedades mentais; Wheeler (2010), que concebe a teoria de Clark como um “funcionalismo estendido”, concorda que esta enfrenta as mesmas dificuldades. Um desses problemas seria, para Menary (2007), a limitação decorrente de uma necessidade de “equivalência de função” entre determinados processos. Afinal, se um computador se torna um artefato devido à função de armazenar dados, e isso é equivalente à função de armazenamento do cérebro, então todos os artefatos precisariam apresentar tal equivalência, quando, na verdade, os processos teriam funcionamentos significativamente diversos. De qualquer forma, o armazenamento dos dados em um computador seria diferente, em muitos sentidos, daquele oferecido pelo organismo biológico.

Para Clark (2008, p. 88), entretanto, aquilo que pode ser considerado um processo cognitivo estendido, ou uma extensão, seria “[...] um padrão amplo de comportamento sistêmico flexível e sensível à informação que sustenta a atribuição de algum estado mental ou atividade cognitiva”. Greif (2015), partindo de um viés evolucionista, adiciona que, tal como o funcionalismo, o local ou a matéria através dos quais determinadas atividades são realizadas são irrelevantes para o debate, posto que a história evolutiva da cognição “incorpora tudo o que acaba por realizar a função em questão”. A dificuldade estaria, ao contrário, em identificar operações que não fossem estendidas.

Viver é claramente uma maneira específica de se relacionar com o ambiente. Também é uma relação interativa e modeladora do ambiente, ao mesmo tempo em que é um sistema cognitivo, em termos da função evoluída da cognição para ajudar um organismo a lidar com um ambiente complexo e em constante mudança. Nesta visão, será difícil encontrar processos cognitivos que não incorporem características do ambiente. Na raiz, toda a cognição é estendida e apenas alguns dos traços cognitivos historicamente mais recentes permitem um notável grau de dissociação. A mente não é estendida para *dentro* do ambiente, mas *sobre* o meio ambiente (Greif, 2015, p.23).

Determinar uma relação direta de equivalência não pareceria, destarte, uma tarefa necessária para a consideração da mente estendida. Tanto a mente quanto o ambiente estariam se inter cruzando e integrando operações “[...] em ambos os lados das fronteiras somáticas” (Greif, 2015, p. 23). A hipótese da mente estendida só pode ser concebida por meio de uma perspectiva que considere mente, corpo e ambiente como componentes de mesma importância e que se constituem como uma mescla cujas funções e papéis se inter põem.

#### **4. A Mente Estendida na esteira do Pós-Humanismo e da Inteligência Artificial**

Dessa forma, se, por um lado, a mente estendida oferece uma proposta de dissolução para o tradicional problema mente-corpo, por outro, conduz a desafios no que tange ao intitulado pós-humanismo e à inteligência artificial. A interseção dos debates entre as ciências cognitivas, a filosofia da mente e a inteligência artificial, portanto, faz-se evidente. Esse amálgama traz consigo, inevitavelmente, aquilo que se denomina “pós-humanismo”. Posto que a inteligência artificial – ou a “*psychological artificial intelligence*”, termo que, conforme Clark (1990), faz menção à estudos de inteligência artificial que se dedicam especificamente à simulação e investigação da cognição – pretende compor sistemas que atuem de forma similar à mente humana, a cognição se torna automaticamente um objeto de estudo relevante para tal campo (Roden, 2015).

Se o pós-humanismo concebe, em todas as suas vertentes, que as perspectivas humanistas elaboradas anteriormente falham em contemplar a totalidade do ser humano, visto que privilegiam determinados caracteres – o que Roden (2015), por exemplo, discorda, ao afirmar que diversas leituras pós-humanistas não parecem considerar as variadas propostas humanistas, empreendendo generalizações que podem distorcer suas próprias argumentações –, o cenário contemporâneo demanda uma nova formulação. Isso significa que, preliminarmente, os modelos de cognição humana se tornam anteriores aos debates sobre o tema e, como sequência, àquilo que é humano ou não.

Nesse sentido, existem divergências naquilo que os teóricos cognitivos entendem como pós-humano. Nayar (2014, p. 11) descreve que o pós-humanismo crítico seria o “[...] descentramento radical do tradicional humano soberano, coerente e autônomo, a fim de demonstrar como o humano já está sempre evoluindo, constituindo e constituído por múltiplas formas de vida e máquinas”. Dessa forma, humanos não podem gozar de um

*status* privilegiado em face de outras espécies, uma vez que sua arquitetura cognitiva não permite um acesso mais ou menos privilegiado ao mundo. Tratar-se-ia, ao contrário, de um organismo em constante união com o ambiente, a tecnologia e outras formas de vida.

Falar do pós-humano é tanto discuti-lo quanto “condição ontológica”, em que a mistura entre humano e não-humano é e será cada vez mais frequente, quanto pensá-lo como resultado de uma demanda por um novo conceito em face das opressões que as discussões anteriores do humanismo geraram. O modo como a ideia de pós-humano e a mente estendida se interseccionam parece radicalizar-se cada vez mais. Se a formulação de Clark é contrária às perspectivas pós-humanistas, compreendendo sua teoria dos humanos enquanto “ciborgues natos” como “argumento da tecnogênese”, é porque a hipótese da mente estendida toma a história dos artefatos como arraigada à história evolutiva da humanidade. A seguinte passagem ilustra o raciocínio:

A transição prometida, ou talvez ameaçada, para um mundo de humanos conectados e dispositivos parcialmente inteligentes é somente mais um passo em um antigo jogo. É um passo que, entretanto, oferece uma oportunidade maravilhosa de pensar mais tempo e mais duramente sobre o que deveria significar ser um humano. Ajuda a ilustrar a condição em que sempre estivemos e mantém um espelho útil para nossos eu's atuais, que na grande maioria não estão conectados. Nós já somos mestres em incorporar profundamente coisas e estruturas não-biológicas em nossas rotinas físicas e cognitivas. Apreciar isso é deixar de acreditar em qualquer futuro pós-humano e resistir à tentação de nos definirmos em brutal oposição aos próprios mundos nos quais tantos de nós atualmente vivemos, amamos e trabalhamos (Clark, 2004, p. 142).

Na tipografia de Clark, Heersmink (2015) problematiza justamente a ausência de um debate acerca das possíveis diferenças entre “artefatos” e “ferramentas” no que tange à utilização humana. A falha da mente estendida parece justamente residir na não citação e enumeração dos artefatos, mas em debater seus aspectos éticos e morais. É possível que, dada a sua ampla aceitação, o uso de artefatos externos seja pensado como convencional e óbvio, especialmente para as ciências interessadas na melhoria das desabilidades cognitivas. A maneira como nos relacionamos com tais ferramentas também seria negligenciada, pois permaneceríamos sem resposta para a pergunta: em que ponto formamos um sistema acoplado ao nos integrarmos com determinadas ferramentas?

A questão anterior não parece ser um foco da hipótese da mente estendida. Clark (1990; 1997; 2004; 2008; 2016), em diversas passagens, afirma que o ser humano é um “ciborgue nato”, ou seja, está intrinsecamente constituindo uma mistura com o ambiente

e as ferramentas nele dispostas. Greif (2015) e Menary (2007) buscam especificar o que seriam artefatos, mas atingem essa mesma conclusão. Segundo Clark (1990), sistemas artificiais de inteligência podem ser limitados para simular determinadas características e habilidades humanas, como interpretação de dados conforme determinados contextos. Porém, o erro em assumir que o computacionalismo pode de fato emular processos cerebrais e, por conseguinte, fenômenos mentais, estaria em outro aspecto. A suposição de uma uniformidade, isto é, a premissa de que somente um sistema seria capaz de realizá-lo, seria o real equívoco, uma vez que a variedade de sistemas permite que alguns se dediquem à simulação de operações básicas (memória espacial, atenção seletiva, reconhecimento de pessoas, entre outros) e outros a processos de alto nível (raciocínio lógico, pensamento consciente etc.).

Estudos relacionados às emoções, como o de Colombetti e Thompson (2007) e Colombetti (2013), empreendem os mesmos esforços que a mente estendida, no sentido de conceber os fenômenos e processos mentais enquanto operações em constante amálgama com o corpo e com o ambiente. De modo similar ao que incentiva Clark (1990), Colombetti (2013, p. 53) aduz que as emoções seriam bem representadas pelos “padrões dinâmicos” ou pela “teoria dos sistemas dinâmicos”. Nesse modelo, proveniente da matemática, ocorreriam “múltiplas e simultâneas interações entre cérebro, corpo e mundo”. Ao aplicá-lo ao domínio das emoções e afetos, a nomenclatura seria “ciência dinâmica afetiva”.

Assim como a mistura que Clark (1990) propõe entre sistemas clássicos e sistemas de distribuição paralela possuiria mecanismos de auto-organização, o modelo descrito por Colombetti (2013) igualmente leva a tal caractere. Para isso, não seria necessário recorrer às representações mentais, uma vez que o objetivo seria de explicar como o sistema apresenta sucesso evolutivo. “Eu defendo que cada episódio emocional deveria ser visto como um padrão dinâmico complexo que foi moldado ao longo do tempo evolutivo e de desenvolvimento” (Colombetti, 2013, p. 57).

## **5. Considerações Finais**

Para além de uma perspectiva que dissolva limites entre o dualismo, o mentalismo e o fisicalismo, é necessário, portanto, que a hipótese da mente estendida não apenas amplie e defina em que consistem os limites da cognição; ela precisa investigar os

artefatos e seus papéis nos processos que permitem perceber o mundo. Heersmink (2015), por exemplo, propõe uma classificação dos artefatos de acordo com três quesitos: (1) consequências para o cérebro, cognição e cultura; (2) o *status* moral de cada artefato; e (3) as relações dos artefatos com a identidade pessoal. Mas, será que conceber como determinadas ferramentas se tornam artefatos da cognição humana, como argumenta Clark (1997, 2004), a partir do momento em que realizam determinada função cognitiva para incorporar estratégias de sobrevivência na vida diária, realmente dissolvem os problemas ontológicos, epistemológicos e éticos que permeiam a natureza dos processos cognitivos?

As interfaces pós-humanistas parecem tornar-se simplesmente vazias ao negar o residual teórico que orbita ao redor da cognição. Roden (2015) afirma que conceber o humano como “ciborgue” não significa dizer que nenhuma medida tecnológica poderá transformá-lo. Ao contrário, o início do uso de símbolos, entre os quais a escrita e a pintura, são exemplos do modo como a arquitetura cognitiva opera para acessar e processar a realidade, ampliando a possibilidade de que outros elementos possam ser observados.

Nesse cenário, a hipótese da mente estendida não é apenas uma espécie de metáfora funcionalista porque, ao contrário, revela aspectos ignorados pela tradição como aqueles que envolvem a percepção e as evidências das ciências cognitivas sobre a ilusão da vida mental e a replicabilidade computacional em seres não-humanos. Contudo, ela ainda parece estar situada em uma fronteira metafísica nebulosa ou, utilizando-nos da afirmação do professor Hofweber (2019, p. 732), “[...] muitas questões metafísicas não têm nada a ver com a forma como representamos a realidade. Mas algumas vezes essa conexão é obtida. Às vezes, pensar em como representamos o mundo pode levar à consequências metafísicas substanciais”.

## **Referências bibliográficas**

Aranyosi, I. [2011]. A New Argument for Mind-Brain Identity. *The British Journal for the Philosophy of Science*, 62(3), 489-517.

Awad, E. *et al.* [2018]. The Moral Machine Experiment. *Nature*, 563, 59-64.

Barack, D.L. [2019]. Mental Machines. *Biology & Philosophy*, 34(63).

Bloor, D. [1970]. Is the Official Theory of Mind Absurd?. *The British Journal for the Philosophy of Science*, 21(2),167-183.

Bundesministerium für Verkehr und digitale Infrastruktur. [2020]. Ethics Commission. Automated and Connected Driving. Disponível em: <[https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission-automated-and-connected-driving.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission-automated-and-connected-driving.pdf?__blob=publicationFile)>. Acessado em 18 de fevereiro de 2020.

Churchland, P. [2013]. *Touching a Nerve: The Self as Brain*. London: W. Norton & Company, 2013.

Clark, A. [1989]. *Microcognition: philosophy, cognitive science and parallel distributed processing*. Cambridge: The MIT Press, 1989.

Clark, A. [1997]. *Being there: putting brain, body and world together again*. 2a ed. Cambridge: The MIT Press, 1997.

Clark, A. [2001]. *Mindware: and introduction to the philosophy of cognitive science*. New York: Oxford University Press, 2001.

Clark, A. [2004]. *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. New York: Oxford University Press, 2004.

Clark, A. [2006]. Language, embodiment and the cognitive niche. *Trends in Cognitive Sciences*, 10(8), 370-374.

Clark, A. [2008]. *Supersizing the Mind: embodiment, action and cognitive extension*. Oxford: OUP, 2008.

Clark, A.; Chalmers, D. [1998]. The Extended Mind. *Analysis*, 58(1), 7-19.

Colombetti, G.; Thompson, E. [2007]. The feeling body: toward an enactive approach to emotion. In: Overton, W.F.; Müller, U.; Newman, J.L. (Eds.). *Developmental Perspectives on Embodiment and Consciousness*. New York: Lawrence Erlbaum Associates, 2007. p. 45-68.

Colombetti, G. [2013]. *The Feeling Body: Affective Science Meets the Enactive Mind*. Cambridge: The MIT Press, 2013.

Greif, H. [2010]. What is the extension of the extended mind? *Synthese*, 194, 4311-4336.

Heersmink, R. [2015]. Extended mind and cognitive enhancement: moral aspects of cognitive artifacts. *Phenomenology and the Cognitive Sciences*, 16, 17-32.

Hochstein, E. [2017]. When does 'Folk Psychology' count as Folk Psychological?. *The British Society for the Philosophy of Science*, 68(4), 1125-1147.

- Hofweber, T. [2019]. Idealism and the Harmony of Thought and Reality. *Mind*, 128(511), 699-734.
- Krellenstein, M. [1995]. Unsolvable problems, visual imagery and explanatory satisfaction. *Journal of Mind and Behavior*, 16, 235–253.
- Nagel, T. [1974]. What is it like to be a bat?. *Philosophical Review*, 83, 435-456.
- Menary, R. [2007]. *Cognitive integration: mind and cognition unbounded*. New York: Palgrave Macmillan, 2007.
- Nayar, P.K. [2014]. *Posthumanism*. Cambridge: Polity Press, 2014.
- Peruzzo Júnior, L. [2019]. Intentionality, Conceptual Content, and Emotions. *Aurora*, 31(54), 833-847.
- Pirenne, M.H. [1950]. Descartes and the Body-Problem in Physiology. *The British Journal for the Philosophy of Science*, I(1), 43-59.
- Roden, D. [2015]. *Posthuman Life: philosophy at the edge of the human*. New York: Routledge, 2015.
- Ryle, G. [2009]. *The Concept of Mind*. London, New York: Routledge, 2009.
- Santana, C. [2019]. Waiting for the Anthropocene. *British Journal for the Philosophy of Science*, 70, 1073-1096.
- Searle, J. [2000]. Minds, Brains and Programs. In: Hofstadter, D.; Dennett, D. (Orgs.). *The Mind's*. New York: Basic Books, 2000.
- Stone, T.; Davies, M. [1993]. Cognitive Neuropsychology and the Philosophy of Mind. *The British Journal for the Philosophy of Science*, 44, 589-622.
- Teixeira, J.F. [2008]. *O que é Inteligência Artificial?*. São Paulo: Brasiliense, 2008.
- Vlerick, M. [2019]. A cognitive perspective on scientific realism. *Philosophical Psychology*, 32(8), 1157-1178.
- Watson, G. [1951]. Apparent Motion and the Mind-Body Problem. *The British Journal for the Philosophy of Science*, II(7), 236-247.
- Wheeler, M. [2010]. In defense of extended functionalism. In: Menary, R. *The Extended Mind*. Cambridge: The MIT Press, 2010. p.245-270.
- Wilks, Y. [1975]. Putnam and Clarke and Mind and Body. *British Journal for the Philosophy of Science*, 26(3), 213-225.