



Dios

MENTE, MUNDO E CORPO: RESPONSABILIDADE E AGENCIAMENTO ESTENDIDO



Murilo Karasinski¹ & Léo Peruzzo Júnior²

Resumo:

Neste artigo pretendemos mostrar como a tese da mente estendida repercute sobre a compreensão das ideias de agenciamento e responsabilidade. Assim, ao considerar que processos cognitivos se estendem além do cérebro e incorporam elementos externos, como ferramentas e tecnologias, discutimos como a responsabilidade estendida requer uma reavaliação dos critérios tradicionais de atribuição de responsabilidade, levando agora em conta também a influência do ambiente na tomada de decisões e nas ações intencionais individuais. Para isso, reconstruímos a tese da mente estendida, para, na sequência, propor uma leitura que considere as implicações dessa nova abordagem às ciências cognitivas.

Palavras-chave:

Mente estendida; Agenciamento; Responsabilidade Estendida; Cognição; Intencionalidade.

Abstract:

In this article, we intend to show how the extended mind thesis impacts the understanding of the ideas of agency and responsibility. Thus, by considering that cognitive processes extended beyond brain and incorporate external elements such as tools and technologies, we discuss how extended responsibility requires a reassessment of traditional criteria for assigning responsibility, now also taking into account the influence of the environment on decision-making and individual intentional actions. To do this, we reconstruct the extended mind thesis and then propose an interpretation that considers the implications of this new approach for cognitive sciences.

Keywords:

Extended Mind; Agency; Extended Responsibility; Cognition; Intentionality.

1 Professor-doutor da Pontifícia Universidade Católica do Paraná – PUCPR. E-mail: k.murilo@pucpr.br
ORCID: <https://orcid.org/0000-0002-6099-6968>

2 Professor-doutor da Pontifícia Universidade Católica do Paraná - PUCPR / FAE. E-mail: leo.junior@pucpr.br
ORCID: <https://orcid.org/0000-0003-3084-5170>

Introdução

Se a tese for aceita, até onde devemos ir? Todos os tipos de quebra-cabeças vêm à mente. E os aldeões amnésicos em Cem Anos de Solidão, que esquecem os nomes de tudo e então penduram etiquetas em todos os lugares? As informações do meu Filofaz contam como parte da minha memória?

Clark & Chalmers (*The Extended Mind*, p.16)

A proposta da mente estendida (*extended mind*), formulada inicialmente por Clark e Chalmers (1998), pretende responder aos desafios da perspectiva dualista, uma vez que indica a conclusão aparentemente radical de que somos uma complexa mistura de nosso corpo, do meio ambiente e de todas as interações empreendidas. Por isso, é uma tese construída sobre a hipótese de que nossos processos cognitivos são *constituídos* por processos tanto derivados do próprio corpo quanto derivados do ambiente e de suas interações.

Entretanto, se assumirmos o tradicional argumento de que a cognição é constituída por ações intencionais, o ambiente desempenha, na teoria da mente estendida, um papel apenas nos casos em que aquilo que é chamado de ‘cognitivo’ pode ser realizado dentro do crânio. Isso significa que as ações intencionais internas podem realizar ações funcionalmente equivalentes fora da cabeça, o que implica que a fronteira entre o que consideramos como processos cognitivos puramente internos e o ambiente externo se torna menos clara. Nesse contexto, o objetivo deste trabalho é explorar as implicações dessa perspectiva para a compreensão da responsabilidade e do agenciamento, uma vez que esses termos estão apoiados sobre a capacidade de tomar as ações como intencionais e conscientes. Por isso, se aceitarmos a premissa de que as ações intencionais podem ter equivalentes funcionais fora da cabeça, isso afeta a nossa compreensão da responsabilidade individual e do próprio agenciamento.

Segundo Clark e Chalmers (1998, p. 7), então, a questão de onde a mente termina e o resto do mundo começa tem recebido duas respostas padrão:

Alguns aceitam os limites da pele e do crânio, e dizem que o que está fora do corpo está fora da mente. Outros são impressionados com argumentos que sugerem que o significado de nossas palavras ‘apenas não está na cabeça’ e sustentam que esse externalismo sobre o significado carrega um externalismo sobre a mente.

É justamente a partir deste confronto que a teoria da mente estendida pretende defender uma terceira posição, particularmente assentada sobre um externalismo baseado no papel ativo do ambiente na condução dos processos cognitivos.

A visão que defendemos aqui é referida por um crescente corpo de pesquisa em Ciência Cognitiva. Em áreas tão diversas quanto a teoria da cognição situada (SUCHMAN, 19870), estudos de robótica do mundo real (BEER, 1989), dinâmicas de abordagens para o desenvolvimento infantil (THELEN e SMITH, 1994), e pesquisas sobre as propriedades cognitivas de grupos de agentes (HUTCHINS, 1995), a cognição é muitas vezes considerada contínua com os processos no ambiente. (...) Alguns acham esse tipo de externalismo intragável. Uma razão pode ser que muitos identificam o cognitivo com o consciente, e parece longe de ser plausível que a consciência se estenda para fora da cabeça nesses casos (CLARK & CHALMERS, 1998, p. 10).

Clark e Chalmers (1998) sustentam que, ao contrário do externalismo passivo de Putnam (1975) e Burge (1998), o organismo humano está ligado ao ambiente numa direção bidirecional, a qual é responsável por criar um sistema acoplado que pode ver visto como um sistema cognitivo por direito próprio. Assim, o papel causal não é desempenhado apenas por um lado do sistema, uma vez que removendo o componente externo, por exemplo, é possível que a competência comportamental perca a sua função. Segundo eles, embora Putnam (1975) e Burge (1998) estejam certos sobre a externalidade do conteúdo, não está claro que tais aspectos externos desempenhem um papel causal ou explicativo na geração da ação (CLARK; CHALMERS, 1998, p. 9). Assim, enquanto para Putnam (1975, p. 151), por exemplo, “o significado determina a extensão, mas pelo fato de a extensão fazer parte do significado”, pois, por um lado, não é o falante por si só que fixa a extensão e, por outro, porque não é o seu estado mental que determina a referência, para Burge (1998, p. 59), o conteúdo de determinadas crenças e estados mentais são “o que uma pessoa crê que significam as suas palavras, como as toma, determina completamente que atitudes pode expressar ao usá-las”, uma vez que o conteúdo de tais eventos ou estados mentais está estritamente limitado a noções que ele compreende.

Seguindo a tipografia do externalismo, a tese da mente estendida procura indicar que objetos externos, assim como a atividade neural, constituem a base dos processos cognitivos e extravasam os limites intracranianos. Isso quer dizer, em outras palavras, que se

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 (como alegamos) parte do processo cognitivo (CLARK & CHALMERS, 1998, p. 8).

Assim, apoiando-se no que foi chamado de *princípio da paridade*, Clark e Chalmers (1998) compreendem que quando um processo cognitivo é substituído por outro e cumpre a mesma função, então, este processo faz parte tanto quando aquele inicial. É justamente a partir do princípio da paridade que os processos cognitivos podem ser *estendidos* para fora do corpo (CARVALHO, 2020, p. 146).

O uso do termo *estendido*, assim como observa Farkas (2011, p. 441), é metafórico, isto é, significa que “estendemos o que conta como um papel funcional ou um perfil disposicional que qualifica um estado como um certo tipo de estado permanente”. Por isso, aqui, falar de “extensão” não é o mesmo que falar de localização espacial, mas do papel que o estado estensional desempenha na vida mental do sujeito. A proposta da mente estendida, por conseguinte, propõe que a base física dos eventos mentais pode se estender além dos limites do nosso corpo orgânico. Considerar a extensão da base física para fora do crânio, por sua vez, exige repensar, conseqüentemente, a localização dos fatores constitutivos dos processos cognitivos, assim como o que significa a ideia de responsabilidade e agenciamento nestes novos termos.

Este artigo pretende mostrar, portanto, que a tese da mente estendida, levada às últimas conseqüências, implica também uma extensão paralela da responsabilidade com os artefatos compartilhados e, ainda, com às agências estendidas como um todo, e não apenas às suas partes humanas. Obviamente, tem-se como resultado, primeiramente, uma nova leitura do individualismo metodológico, para o qual toda ação social deveria ser explicada em termos de motivos e disposições e, posteriormente, o argumento de que uma ação é a combinação de elementos humanos e não-humanos, estes últimos agora acoplados. É justamente este último argumento que indica a necessidade de reavaliar os limites tradicionais da responsabilidade e atribuí-la não apenas aos agentes humanos, mas também aos artefatos e agências estendidas que estão envolvidos na ação.

1. Artefatos não-biológicos e acoplamento: novos andaimes cognitivos

Os exemplos de Otto e Inga, apresentados por Clark e Chalmers (1998, p. 12), são utilizados para ilustrar a ideia de que o uso de artefatos externos pode ser integrada de tal forma que se tornem parte essencial do processo cognitivo. No exemplo de Otto, um indivíduo que sofre de uma deficiência de memória (doença de Alzheimer) utiliza um bloco de notas para compensar sua limitação. Ele anota informações importantes no bloco de notas, como compromissos e endereços, e consulta essas anotações sempre que necessário. Nesse caso, o bloco de notas funciona como uma extensão da memória de Otto, permitindo que ele acesse informações de forma rápida e eficiente. Isso não seria diferente de um indivíduo que armazena uma série de notas em seu celular ou que opta por um sistema *iCloud* para armazenar dados, como documentos e fotos, os quais podem ser baixados em outro momento.

Já o exemplo de Inga envolve uma situação semelhante, mas com uma diferença crucial. Inga não possui deficiência de memória, mas opta por utilizar suas crenças para caminhar até o Museu de Arte Moderna, que está localizado na 53rd Street. Sua crença estava em algum lugar na memória, esperando ser acessada. Segundo Clark e Chalmers (1998, p. 14), seria inutilmente complexo explicar as ações de Inga em termos de crenças sobre sua memória. Por isso, “o bloco é uma constante para Otto, da mesma forma que a memória é uma constante para Inga”. Nos dois casos, os artefatos (o bloco de notas de Otto e as crenças de Inga) desempenham um papel fundamental na cognição, pois eles são utilizados como recursos confiáveis para auxiliar nas tarefas cognitivas, assim como para acessar informações importantes e orientar ações específicas.

Clark e Chalmers (1998) argumentam, então, que explicar as ações de Inga em termos de crenças seria excessivamente complexo. Em vez disso, eles sugerem que a mente estendida pode ser compreendida por meio da interação entre os elementos cognitivos internos (como as crenças de Inga) e os artefatos externos (como o bloco de notas de Otto). Os autores ainda procuram atacar o argumento de que, em alguns casos, a informação externa poderia não estar disponível, o que prejudicaria o exemplo de Otto e endossaria a posição de que as crenças de Inga sempre estarão disponíveis em sua memória:

Pode-se temer que o acesso de Otto ao seu caderno *de fato* venha e vá. Ele toma banho sem o caderno, por exemplo, e não consegue lê-lo quando está escuro. Certamente sua crença não pode ir e vir tão facilmente? Poderíamos contornar esse problema redescrivendo a situação, mas, em qualquer caso, uma desconexão temporária ocasional não ameaça nossa reivindicação. Afinal, quando Inga

está dormindo, ou quando está embriagada, não dizemos que a sua crença desaparece. O que realmente conta é que a informação está facilmente disponível quando o sujeito precisa dela, e essa restrição é satisfeita igualmente nos dois casos. Se o caderno de Otto muitas vezes não estivesse disponível para ele nos momentos em que as informações nele contidas seriam úteis, poderia haver um problema, pois a informação não seria capaz de desempenhar o papel de guia de ação que é central para a crença; mas se estiver facilmente disponível na maioria das situações relevantes, a crença não está em perigo (CLARK & CHALMERS, 1998, p. 15).

Parece claro, portanto, que essa perspectiva ressalta que a cognição estendida não exige uma disponibilidade perfeita e ininterrupta dos artefatos externos, mas uma disponibilidade prática que permita uma interação efetiva entre o indivíduo e os recursos externos quando necessário. Embora existam momentos em que o acesso ao caderno de Otto pode ser temporariamente limitado, como durante o banho ou em situações de baixa luminosidade, os autores argumentam que essas interrupções não invalidam a concepção da mente estendida. O critério essencial é que a informação esteja facilmente disponível na maioria das situações relevantes, onde a crença precisa ser utilizada como guia para a ação. Por isso, essa abordagem reconhece que os processos cognitivos podem se estender além dos limites do cérebro individual e incorporar elementos externos ao sujeito.

O interessante dessa abordagem é que ela prioriza os artefatos externos como novos andaimes cognitivos, reconhecendo o papel de tais artefatos como extensões funcionais e complementares dos processos funcionais. Ao mesmo tempo, ao afirmar que os processos cognitivos se estendem para além dos limites do cérebro individual, ela desafia a visão tradicional de que a cognição é estritamente confinada ao domínio interno da mente. O fato é que, sendo a cognição fortemente influenciada e moldada pela interação entre nossos processos internos e os recursos e ferramentas disponíveis externamente, temos como uma das consequências diretas a própria mudança nos conceitos de *agenciamento* e *responsabilização*. Afinal, não é suficiente reconhecer que a mente não é um fenômeno interno, mas uma entidade em constante interação com o mundo ao seu redor. Precisamos determinar quando uma crença é suficientemente reconhecida como verdadeira e em que medida deveríamos dissolver a responsabilidade entre o agente e o próprio artefato que está sendo usado.

Otto poderia reconhecer no bloco de notas uma série de informações que lhe permitiriam transitar entre os ambientes do Museu. Seria capaz de saber onde estão cada uma das coleções e dos quadros que estão sendo expostos, porém o bloco de notas é uma ferramenta meramente passiva que apenas facilita o acesso às informações e que não tem

uma capacidade cognitiva inerente. Otto poderia confiar, então, que as informações contidas no bloco de notas sejam adequadamente corretas, quando na verdade elas são falhas e conduzem a raciocínios equivocados. Se estou procurando recordar algo que não sei, como posso confiar que um processo cognitivo estendido seja capaz de tornar novamente emergente aquela informação?

Como se argumenta, não é suficiente apontar que o cérebro é um poderoso processador de informações, pois diante dos limites de sua capacidade de lembrar, calcular, navegar, planejar e aprender, faz uso de artefatos que nos sustentam na execução de tarefas cognitivas (HEERSMINK, 2017). Por isso, a capacidade de alguém agir intencionalmente no mundo, tomar decisões, fazer escolhas e executar ações com base em crenças, desejos e intenções implica um sentido de controle e autodeterminação sobre as próprias ações. O uso dos artefatos não-biológicos, como canetas, *smartphones* ou qualquer outro recurso tecnológico, então, permite que novos andaimes cognitivos possam ser construídos, estendendo as capacidades mentais além dos limites naturais do cérebro. Esses andaimes cognitivos funcionam como suportes estendidos que complementam e aprimoram nossas habilidades cognitivas, possibilitando a resolução de problemas mais complexos que, de outra forma, seriam inacessíveis para o cérebro isolado.

É provável que, em alguns casos, os artefatos forneçam recursos adicionais para moldar e influenciar como pensamos e nos comportamos, mas isso não significa que tais artefatos possam substituir a capacidade de agirmos intencionalmente. Em vez disso, os artefatos são usados como ferramentas ou auxílios para facilitar a expressão da agência. Há várias razões que poderiam explicar porque o cérebro biológico teria evoluído e amadurecido com tais capacidades para manipular o ambiente externo. Uma dessas razões seria a tentativa do cérebro em parasitar o ambiente para reduzir a carga de memória e até mesmo transformar a natureza dos próprios problemas computacionais (CLARK & CHALMERS, 1998). Foi essa habilidade de distribuir-se pelo ambiente que permitiu potencializar nossa capacidade de agir intencionalmente no mundo e, conseqüentemente, incorporar o ambiente em nossos processos cognitivos.

Por isso, a leitura tradicional de que o agenciamento e a responsabilidade possam situar-se apenas no poder de agir de acordo com suas próprias intenções ou na capacidade de decidir e exercer controle sobre as próprias ações e comportamentos precisa ser reconsiderada. A concepção tradicional de agenciamento e responsabilidade costuma focar a ideia de que o indivíduo é totalmente autônomo e livre para tomar decisões, como se fosse uma entidade isolada do ambiente. No entanto, essa abordagem não é capaz de determinar o papel dos artefatos não-biológicos nos processos cognitivos e nem explicar as interações que mantêm durante o seu acoplamento.

2. Agenciamento e externismo ativo: acoplando os artefatos à responsabilidade

A teoria da mente estendida, conforme concebem Clark e Chalmers (1998), se movimenta sobre o que os autores têm denominado de “externismo ativo”, isto é, um tipo de interação que os indivíduos mantêm com o ambiente e que envolve as suas ações (CARVALHO, 2020; ALONSO; RAMOS, 2022; VIERKANT, 2022). Enquanto um externismo de conteúdo apontaria para o histórico de interações causais e para o modo como o ambiente contribui para a determinação do conteúdo de nossos estados mentais, o externismo ativo compreende que a mente está imersa em um determinado contexto e profundamente entrelaçada com ele. Por isso, enquanto as teses tradicionais das Ciências Cognitivas, dentre as quais o cognitivismo, optaram por perspectivas descorporificadas, a posição da mente estendida se afirma na dinâmica percepção-ação, um enfoque que reconhece a interação constante entre o sujeito cognitivo e seu entorno.

O externismo ativo, então, reconceitualiza a mente não como um ente isolado, mas como um sistema que se aproveita ativamente do ambiente para facilitar processos de pensamento, resolução de problemas e tomada de decisões. Isso significa, portanto, que o externismo ativo vai além da mera influência do ambiente, enfatizando a coparticipação ativa do ambiente em nossos processos cognitivos. Nessa perspectiva, a cognição opera em um ciclo contínuo de percepção e ação, onde as fronteiras tradicionais entre a mente e o mundo se tornam mais fluidas, permitindo que utilizemos ferramentas, tecnologias e outros elementos do ambiente como extensões das nossas capacidades cognitivas. Em outras palavras, o externismo ativo equivale a dizer que alguns de nossos estados mentais são constituídos por elementos que estão “fora de nossa cabeça”.

É verdade, porém, que assumindo a tese da mente estendida precisamos evocar, por um lado, uma visão funcionalista mínima e, por outro, a existência de um banco de crenças estocadas na própria memória. Em relação ao funcionalismo, os processos mentais podem ser compreendidos como a função que desempenham em um sistema cognitivo. A função que tais processos exercem no próprio sistema é suficiente para considerá-los como estados capazes de realizar cognição (CLARK & CHALMERS, 1998, p. 7-8). O fato é que essa posição desobriga a necessidade de que tais estados mentais sejam realizados por um cérebro, pois o que está em jogo é a forma como isso é capaz de produzir conteúdo e comportamento.

Já em relação ao banco de crenças estocados na própria memória, a questão parece ser um pouco mais profunda. É difícil justificar boa parte de nossas ações sem referência

a um conjunto de crenças e desejos que não nos ocorrem o tempo todo, mas que estão ali quando outras suposições passam a ser consideradas, como é o caso das emoções (PERUZZO JÚNIOR, 2019). Por exemplo, se posso confrontar a tarefa com uma parte do mundo e isso funciona adequadamente bem, então, deveríamos reconhecer como parte de um processo essa parte do mundo.

Assim, no contexto do externismo ativo, as crenças armazenadas em nossa memória também são afetadas por essa interação. Para tal, nossas crenças não são formadas em um vácuo, mas são moldadas pelas experiências e interações que temos com processos que poderiam ocorrer externamente. Isso significa que o externo não é apenas um fator que influencia nossa cognição, pois ele desempenha um papel ativo na forja das estruturas cognitivas. Segundo Clark (1997, p. 214), “as complexidades da dinâmica usuário-artefato convidam à reflexão sobre um tópico mais geral: como conceber a fronteira entre o sistema inteligente e o mundo. Esta fronteira (...) parece ser um pouco mais plástica do que se supunha anteriormente”.

Assumindo essa infiltração da mente no mundo, as Ciências Cognitivas precisam lidar com a própria ideia de artefatos e, com isso, redesenhar o argumento de que o agenciamento possa ser pensado apenas como a capacidade dos seres racionais de realizarem ações intencionais e serem responsáveis por suas escolhas. Por isso, se considerarmos a teoria da mente estendida, há uma reviravolta significativa no conceito de agenciamento (*agency*): o agenciamento não é visto apenas como uma propriedade do cérebro individual, mas também como algo que pode se estender para além dos limites corporais. Tendo em vista essa leitura, precisamos considerar a maneira como o ambiente funciona para complementar e estender nossas capacidades cognitivas e de tomada de decisão. Isso implica, conseqüentemente, repensar o argumento de que a responsabilidade deva se concentrar exclusivamente nos processos individuais internos.

De qualquer forma, aceitar que o compartilhamento da responsabilidade com artefatos não humanos possa diluir a responsabilidade moral dos humanos amplia a própria ideia de agenciamento, levantando dúvidas sobre como determinar quem é o responsável quando as decisões são feitas em colaboração com as tecnologias. Assim, à medida que os artefatos desempenham um papel mais central, precisamos desenvolver novas normas e regulamentações para lidar com as implicações éticas da responsabilidade compartilhada. Talvez aqui Clark (1997, p. 215) tenha razão ao afirmar que existe, afinal, uma dificuldade bastante geral em traçar uma linha firme entre um usuário e uma ferramenta: “Uma pedra segurada na mão e usada para quebrar uma noz é claramente uma ferramenta. Mas se um

pássaro deixar cair uma noz do ar para que ela se quebre ao entrar em contato com o solo, o solo é uma ferramenta?”. Por isso, não se trata de reduzir a responsabilidade e o agenciamento à ação situada, mas repensar a forma como nós humanos devemos compreender os próprios artefatos, considerando que não podemos ser exclusivamente responsáveis por isso.

3. Repensando a responsabilidade: a ideia de ação situada

Para David Hume (2006, p. 32), todos os raciocínios acerca de questões factuais parecem fundamentar-se na relação de causa e efeito. Apenas por meio dessa relação é possível ultrapassar a evidência da memória e dos sentidos. Um observador que encontrasse um relógio ou qualquer outra máquina em uma ilha deserta concluiria que em algum momento houve um ser humano naquela ilha. Segundo Hume (2006, p. 33), as causas e os efeitos não podem ser descobertos pela razão, mas sim pela experiência. “Ninguém imagina que a explosão da pólvora ou a atração de um ímã poderiam descobrir-se por meios de argumentos *a priori*” (Hume, 2006, p. 33).

Com efeito, quando se aborda o tema da responsabilidade é crucial considerar, por exemplo, a análise que o Direito Penal faz desse assunto. Especificamente no âmbito jurídico, sobretudo no campo criminal, é fundamental compreender se alguém pode ser responsabilizado por um crime, levando em conta a existência do nexos causal que conecta a conduta (causa) ao resultado naturalístico (alteração no mundo externo).

No contexto brasileiro, por exemplo, o artigo 13 do Código Penal define a causa como a ação ou omissão sem a qual o resultado não teria ocorrido, adotando assim o que é conhecido como “teoria da equivalência dos antecedentes causais” (*conditio sine qua non*). Essa teoria sustenta que a conduta deve ser considerada a causa quando o resultado não teria ocorrido na ausência dessa conduta.

Deve-se somar à teoria da equivalência dos antecedentes causais o procedimento hipotético de eliminação de Thyrén, que reputa como causa todo evento que, ao ser mentalmente removido, resultaria no desaparecimento do resultado conforme ocorreu ou quando ocorreu. Por exemplo, no caso em que um indivíduo disparasse uma arma contra outro (conduta) resultando na morte da vítima (resultado) devido ao disparo (nexo causal), o método de eliminação hipotética de Thyrén asseguraria que, se o primeiro indivíduo não tivesse efetuado o disparo, o resultado de morte não teria ocorrido.

Além disso, para evitar uma regressão infinita, é fundamental que se estabeleçam critérios claros de demarcação. No caso de um fabricante de armas, para mencionar um exemplo recorrente, poderia ser discutida a responsabilidade pela produção de uma arma que eventualmente fosse utilizada para cometer um crime. Para que haja uma atribuição de responsabilidade, porém, é preciso considerar se o fabricante agiu com dolo ou culpa (responsabilidade subjetiva), ou se ele criou um risco relevante e juridicamente proibido (imputação objetiva).

Nesse sentido, retornando-se à questão do agenciamento estendido, John Protevi (2008) investigou o papel dos afetos, como raiva e pânico, e dos reflexos pré-cognitivos no contexto dos combates e do ato de matar. Segundo Protevi (2008), apenas 2% da população seria capaz de matar de forma fria, olho no olho, consciente do que está fazendo, sendo esses indivíduos categorizados como de baixa afetividade ou sociopatas em busca de estímulos. Protevi (2008) argumentou que a dicotomia “lutar ou correr” seria erroneamente extrapolada de conflitos entre espécies, nos quais é mais provável que a fuga ocorresse antes da luta efetiva. Embora seja verdade que em algumas espécies territoriais, como os leões, um novo macho alfa vitorioso possa vir a matar a prole do adversário derrotado, a inibição de atos violentos dentro da espécie seria a regra. Um exemplo que ilustra essa restrição poderia ser observado nos conflitos e assassinatos entre chimpanzés, nos quais as emboscadas frequentemente envolvem não apenas dois, mas muitas vezes sete ou oito animais atacando uma única vítima isolada.

No contexto da Segunda Guerra Mundial, apesar dos avanços no treinamento em equipe e no comando, que aumentaram a capacidade de envolvimento em combates letais, Protevi (2008) indicou que a taxa de disparos à queima-roupa das tropas de infantaria americana foi de apenas 15% a 20% do número total de tiros efetuados durante a guerra. Isso sugeriu que grande parte dos disparos realizados em combate foi mais uma demonstração de presença do que uma efetiva luta. Nesse âmbito, conforme observado por Protevi (2008), a maioria dos soldados não conseguiu realizar assassinatos com frieza, necessitando matar em um estado de dessubjetivação, ou seja, movidos por reflexos, fúria e pânico. Diante disso, surge a questão levantada por Protevi (2008): quem é o responsável pelos atos de violência quando os reflexos, a fúria e o pânico são ativados?

Para abordar essa questão, é possível encontrar argumentos no pensamento de Allan Hanson. Segundo Hanson (2008), os teóricos sociais estão cada vez mais conscientes de que todas as ações dos seres humanos, ou até mesmo as possíveis ações, não ocorrem isoladamente. Em vez disso, os seres humanos sempre atuam em conjunto, seja

com outras pessoas, ferramentas e objetos, animais, computadores ou, o mais comum, por intermédio de uma combinação de todos esses elementos. De acordo com Hanson (2008), seria impossível, por exemplo, realizar uma busca por um livro no catálogo de uma biblioteca *on-line* sem a participação ativa de um computador. Da mesma forma, percorrer grandes distâncias seria impensável sem o uso de um carro, trem ou avião.

Com base nesse raciocínio, Hanson (2008) argumentou que a responsabilidade moral deveria ser atribuída não apenas aos seres humanos, mas também às agências estendidas que englobariam elementos não-humanos. Esse enfoque é fundamentado no pensamento pós-modernista, que destaca que as ações são realizadas e devem ser explicadas em termos de combinações interrelacionadas entre os humanos e os não-humanos.

Com efeito, Hanson (2008) sustentou que a visão tradicionalmente humanista do indivíduo como um sujeito, originada no Renascimento e consolidada no início da modernidade, teria se tornado obsoleta como estratégia para explicar as ações morais, o que demandaria, por seu turno, que fosse substituída por explicações originadas nessa agência estendida, a qual é “fundamentalmente diferente do indivíduo humano e, portanto, as formas tradicionais de pensar sobre a responsabilidade moral dos seres humanos não se aplicam” (Hanson, 2008).

De acordo com Hanson (2008), a agência estendida se caracterizaria por ser um conceito aberto e ilimitado, permitindo a inclusão progressiva de mais participantes. Isso poderia acarretar uma reavaliação da responsabilidade, inclusive no que diz respeito à questão da regressão infinita como critério para a demarcação do Direito Penal, mencionada anteriormente, além de trazer novos elementos para a discussão da causalidade, na forma clássica proposta por Hume. Nesse contexto, a agência responsável por um homicídio em um acidente de trânsito poderia abranger diversos elementos, como as condições da via e as condições climáticas, entre outros, dependendo das circunstâncias específicas. Segundo Hanson (2008), existiria um ponto de retornos decrescentes, em que a participação de certos elementos se tornaria menos relevante (por exemplo, a cor da roupa da vítima, que pode influenciar a visibilidade do motorista, mas a diferença entre tons de marrom ou azul escuro seria de menor importância). Além disso, haveria fatores constantes e universais, como o efeito da gravidade, que deveriam ser desconsiderados.

Para Hanson (2008), uma perspectiva moral fundamentada em uma ética deontológica, na qual o pensamento de Kant se destaca, tenderia a restringir a responsabilidade moral apenas aos agentes capazes de agir intencionalmente, ou seja, aos seres humanos, o que seria chamado de *individualismo moral*. No entanto, Hanson (2008) pondera a noção

de uma agência estendida que incluísse tanto elementos humanos quanto não humanos, atuando de forma conjunta e intencional. Para ilustrar essa ideia, é apresentado o cenário hipotético de uma busca por uma pessoa perdida em uma região selvagem. Nesse caso, a operação de busca envolveria uma variedade de componentes, como cachorros farejadores, carregadores humanos, helicópteros com suas tripulações, equipes em solo realizando buscas, bem como o pessoal no centro de comando coordenando todas as atividades e os equipamentos de comunicação utilizados para manter a comunicação entre os envolvidos. Essa configuração completa constituiria a agência estendida responsável pela busca.

Por outro lado, o conceito de agência estendida apresentado por Hanson (2008) é debatido por aqueles que defendem a perspectiva do individualismo moral. De acordo com esta visão, apenas os seres humanos podem ser atribuídos com intenções e, portanto, somente eles poderiam ser considerados moralmente responsáveis por suas ações. Consequentemente, os cachorros, helicópteros e equipamentos de comunicação não seriam considerados moralmente responsáveis por estarem envolvidos na busca pela pessoa perdida. Para Hanson (2008), no entanto, a dificuldade com essa forma de pensar é que ela partiria do pressuposto de que os elementos não humanos seriam apenas ferramentas utilizadas por seres humanos estáveis e autônomos para atingir seus objetivos. Essa visão negligenciaria o fato de que os seres humanos seriam afetados e modificados pelo uso de determinadas tecnologias, como evidenciado quando se percebe que um homem com uma arma é diferente do mesmo homem desarmado; ou quando uma pessoa se transforma quando está atrás do volante de um carro.

Questão complementar no debate sobre a agência estendida diz respeito ao tratamento de recompensas e punições, que normalmente se aplicariam somente aos seres humanos. Carros, armas, computadores ou qualquer outra entidade não-humana não sofreriam consequências positivas ou negativas decorrentes de um determinado comportamento. Por exemplo, não faria sentido encarcerar uma arma envolvida em um crime. Contudo, segundo Hanson (2008), se o propósito de aplicar uma sanção seria incentivar certos tipos de ações em detrimento de outros, o foco deveria recair sobre as agências estendidas que produzem essas ações. Nesse contexto, poder-se-ia argumentar – no sentido de responsabilizar as agências estendidas – que quando uma criança se comporta mal, usando um brinquedo de forma violenta ou se recusando a compartilhá-lo, uma punição típica seria retirar o seu brinquedo. Hanson (2008) destaca que, ao liberar o *hacker* de computador Kevin Mitnick da prisão, a decisão judicial o proibiu de utilizar equipamentos por três anos como parte de sua punição. “Assim, é razoável entender que a punição é dirigida contra o relacionamento ofensivo como um todo, e não apenas contra uma parte

específica dele” (HANSON, 2008), de modo que a atuação da agência estendida seria eliminada, evitando, além disso, o surgimento de outras entidades similares, que, embora pudessem apresentar variações em seus componentes, ainda representariam ameaças capazes de causar danos do mesmo tipo ou semelhantes.

Outro aspecto que merece destaque é a concepção de inteligência desenvolvida por Santaella (2003), com base no pensamento de C. S. Peirce. Segundo Santaella (2023, p. 100), a presença de inteligência está relacionada à capacidade de aprender, crescer ou evoluir, independentemente do quão primitiva essa ação possa parecer. Como resultado, Santaella (2023, p. 100) sugere que o pensamento não se limitaria ao cérebro humano, uma vez que já existiriam estudos que fundamentam a ideia de que florestas e plantas também possuiriam capacidade de pensamento. De fato, esses ecossistemas poderiam cuidar de seus vizinhos doentes, alertar suas companheiras sobre perigos através de sinais elétricos transmitidos por redes de fungos e até mesmo fornecer uma solução açucarada para as raízes de árvores caídas há séculos, mantendo-as vivas.

De acordo com Santaella (2023, p. 108), é incontestável que programas de computador e algoritmos de IA demonstrariam algum nível de inteligência, desde que não fossem comparados exclusivamente com o padrão da inteligência humana:

Uma discussão mais clara sobre a presença ou não da inteligência na IA tem sido atrapalhada pelo fato de que a IA é alimentada por dados humanos, esses dados são selecionados por humanos, o design é modelado por humanos, os vieses são apontados por humanos etc. Ou seja, a intervenção humana é extensa demais para que, aparentemente, se possa reconhecer a presença da inteligência nas máquinas. Essa falta de reconhecimento poderia ser verdadeira se um fator muito relevante não fosse deixado para trás, a saber, os algoritmos de IA aprendem. Não é do acaso que surgiram os nomes de “aprendizagem de máquina” e, mais ainda, “aprendizagem profunda”. Onde existe aprendizagem, agenciamentos inteligentes entram em cena. Aprender implica um maior ou menor grau de autonomia transformadora. Essa autonomia, evidentemente em grau bem inferior à autonomia humana, está sendo assumida pela máquina (SANTAELLA, 2023, p. 108-9).

Hanson (2008) defende um raciocínio semelhante, argumentando que está cada vez mais evidente que os seres humanos não ocupariam um lugar especial no mundo, dada a sua conexão com outras formas de vida. Portanto, qualquer teoria que restringisse a responsabilidade moral exclusivamente aos indivíduos humanos estaria completamente desatualizada devido à abundância de interações, inter-relacionamentos e interdependência entre humanos e não-humanos.

Por outro lado, de acordo com Ronald Giere (2008), a ideia de distribuir a responsabilidade moral entre seres humanos e os artefatos que produzem resultados moralmente relevantes no mundo não se sustentaria. Por exemplo, helicópteros apenas executariam suas funções designadas, voando em determinadas velocidades e pairando quando necessário. Os elogios morais, conforme Giere (2008), deveriam ser direcionados apenas aos pilotos que, por escolha própria, decidiriam voar na altura das árvores para arriscar suas vidas e salvar pessoas perdidas. Compartilhar a responsabilidade moral com elementos não humanos, segundo Giere (2008), atenuaria a responsabilidade moral dos seres humanos.

Contudo, essa perspectiva não é consensual. Segundo Cash (2010), os indivíduos seriam apenas um dos possíveis centros de cognição, ação e responsabilidade. Assim, o aspecto mais relevante da investigação sobre responsabilidade e cognição estaria em enfatizar a responsabilidade coletiva pelas ferramentas cognitivas e práticas normativas que sustentam a agência estendida. Consoante Cash (2010), isso implicaria, frequentemente, mas não sempre, utilizar as premissas do individualismo moral.

Em nossa leitura, seria crucial enfatizar a responsabilidade coletiva em compreender e abordar de maneira mais direta os diferentes aspectos do contexto em que as pessoas vivem, pensam e agem. Essa abordagem possui respaldo nas considerações de Hutchins (1995, p. xiii), que destacava a importância de se considerar a atividade cognitiva a partir de um contexto, entendendo-o como um processo dinâmico mais amplo, e não como um conjunto estático de condições do ambiente. Concordando com a visão de Hayles (1999, p. 289), que argumentava que a sofisticação cognitiva dos seres humanos modernos se deveria à construção de ambientes mais inteligentes para trabalhar, torna-se evidente a necessidade de efetuarmos uma nova compreensão sobre a responsabilidade e o agenciamento estendido, uma vez que a cognição tem como característica misturar-se ao ambiente para encontrar formas de suporte e otimização.

Considerações finais

O avanço constante de tecnologias e ferramentas tem possibilitado a modificação do ambiente ao qual os seres humanos estão expostos através de artefatos que, cada vez mais, são incorporados para aprimorar a cognição humana. No entanto, muitos dos conceitos morais e jurídicos ainda em uso são baseados em construções antiquadas que há muito tempo deveriam ter sido reavaliadas e atualizadas. Nesse sentido, o agenciamento

to estendido, por um lado, é uma perspectiva que traz novas visões sobre o mundo e a responsabilidade humana. É crucial aprofundar esse conceito para as ciências que se desenvolverão ao longo do século XXI, pois ele oferece uma compreensão mais abrangente de como as tecnologias e ferramentas estão integradas à experiência humana e como elas afetam a noção de mente, mundo e corpo.

Por outro lado, ao estendermos nossa agência mental para artefatos não humanos, podemos estar diluindo o senso de responsabilidade moral, uma vez que a própria ideia de responsabilidade está ligada à intenção e à causalidade. Assim, deveríamos estabelecer, antes de tudo, em que medida um artefato não humano contribui para uma decisão ou ação para, na sequência, vincular o possível grau de controle que alguém possa ter sobre uma situação ou evento. À medida que assumimos essa nova leitura sobre o agenciamento e a responsabilidade, o desafio reside em equilibrar a expansão de nossa capacidade cognitiva em relação à função desempenhada pelos artefatos tecnológicos a fim de que a colaboração entre humanos e tecnologia não apenas amplie a justificação epistêmica de nossas ações, mas também preserve os fundamentos da responsabilidade individual e coletiva.

Referências

ALONSO, Bernardo Gonçalves & RAMOS, Ronaldo de Oliveira. A marca do Cognitivo e Cognição 4E. **Princípios: Revista de Filosofia**, Natal, v.29, n.58, jan./abr. 2022, p.24-48.

BURGE, Tyler. Individualism and the Mental. In: LUDLOW, P & MARTIN, N. (Eds.). **Externalism and self-knowledge**. Stanford: CSLI Publications, 1998.

CARVALHO, Eros Moreira de. A tese da mente estendida à luz do externismo ativo: como tornar Otto responsivo a razões? **Trans/Form/Ação**, 43(3), Jul-Sep. 2020, p.143-166.

CARTER, Adam J.; CLARK, Andy; KALLESTRUP, Jesper; PALERMOS, Orestis S. & PRITCHARD, Duncan (Eds.). **Extended Epistemology**. Oxford: Oxford University Press, 2018.

CASH, Mason. Extended cognition, personal responsibility, and relational autonomy. **Phenomenology and the Cognitive Sciences**, 9 (4):645-671, 2010.

CLARK, Andy. **Being there: putting brain, body, and world together again**. Cambridge: The MIT Press, 1997.

CLARK, Andy & CHALMERS, David. The Extended Mind. **Analysis**, v. 58, n.1, jan. 1998, p.7-19.

FARKAS, Katalin. Two versions of the Extended Mind Thesis. **Philosophia**, v.40, 2012, p.435-7.

GIERE, Ronald. Human moral responsibility is moral responsibility enough: A reply to F. Allan Hanson. **Phenomenology and the Cognitive Sciences**, 7, 2008, p. 425-7.

HANSON, F. Allan. The anachronism of moral individualism and the responsibility of extended agency. **Phenomenology and the Cognitive Sciences**, 7, 2008, p.415-24.

HAYLES, N. Katherine. **How we became posthuman: virtual bodies in cybernetics, literature, and informatics**. Chicago: University of Chicago Press, 1999.

HEERSMINK, Richard. Extended Mind and Cognitive Enhancement: Moral Aspects of Cognitive Artifacts. **Phenomenology and the Cognitive Sciences**, 16(1), 2017, p.17-32.

HUTCHINS, Edwin. **Cognition in the wild**. Cambridge: MIT Press, 1995.

HUME, David. **Investigações sobre o entendimento humano**. Tradução André Campos. São Paulo: Escala Educacional, 2006.

PERUZZO Júnior, Léo. Intentionality, Conceptual Content, and Emotions. **Revista de Filosofia Aurora**, Curitiba, v. 31, n. 54, set./dez. 2019, p.833-47.

PERUZZO JÚNIOR, Léo; CANDIOTTO, Kleber & KARASINSKI, Murilo (Orgs.). **Tendências contemporâneas de filosofia da mente e ciências cognitivas**. Curitiba: PUCPress, 2023.

PROTEVI, John. Affect, agency and responsibility: The act of killing in the age of cyborgs. **Phenomenology and the Cognitive Sciences**, 7, 2008, p. 405-13.

PUTNAM, Hilary. It ain't necessarily so. In: Putnam, Hilary. **Mathematics, matter and method**: philosophical papers. New York: Cambridge University Press, v. 1, 1975.

SANTAELLA, Lucia. **A inteligência artificial é inteligente?** São Paulo: Edições 70, 2023.

VIERKANT, Tillmann. **The tinkering mind**: agency, cognition, and the extended mind. Oxford: Oxford University Press, 2022.



