



REPRESENTAÇÃO E COGNIÇÃO SITUADA: UMA PROPOSTA CONCILIADORA PARA AS GUERRAS REPRESENTACIONAIS



Carlos Henrique Barth (Doutorando, PPG-Filosofia, UFMG)

Felipe Nogueira de Carvalho (Professor e Pesquisador, UFLA)

Resumo:

Abordagens pós-cognitivistas mais recentes têm lançado duras críticas à noção de representação mental, procurando ao invés disso pensar a mente e a cognição em termos de ações corporificadas do organismo em seu meio. Embora concordemos com essa concepção, não está claro que ela implique necessariamente a rejeição de qualquer tipo de vocabulário representacional. O objetivo deste artigo é argumentar que representações podem nos comprar uma dimensão explicativa adicional não disponível por outros meios e sugerir que, ao menos em alguns casos, elas podem participar da explicação de performances ou capacidades cognitivas. A noção de representação apresentada, como deixaremos claro ao longo do artigo, não viola os preceitos metodológicos mais caros à cognição 4E, em geral, e ao enativismo, em particular, podendo, portanto, ser utilizada como uma ferramenta teórica útil em investigações sobre a natureza corporificada e situada da mente.

Abstract:

Recent post-cognitivist approaches have raised sharp criticisms against the notion of mental representation, proposing instead to think of the mind and cognition in terms of embodied actions of an organism in its environment. While we agree with this conception, it is not clear that it necessarily implies the rejection of any kind of representational vocabulary. The aim of this paper is to argue that representations afford an additional explanatory dimension that's unavailable through other means, and to suggest that, in at least some cases, they may participate in the explanation of cognitive performances or capacities. The presented notion of representation, as we will make clear throughout the paper, does not violate the methodological precepts most dear to 4E cognition in general and enactivism in particular, and can therefore be used as a useful theoretical tool in investigations about the embodied and situated nature of the mind.

Palavras-chave:

Representação mental; cognição situada; cognição 4E; enativismo.

Keywords:

Mental representation; situated cognition; 4E cognition; enactivism.

Introdução: breve relato das guerras representacionais

Embora a noção de ‘representação’ tenha desempenhado um papel fundamental no desenvolvimento das ciências cognitivas a partir dos anos 50 e 60, abordagens pós-cognitivistas mais recentes – notadamente o enativismo – têm lançado duras críticas a esta noção, redesenhando o estudo da mente de forma a dispensar qualquer tipo de vocabulário representacional. Assim, ao invés de pensar a cognição em termos de computações sobre estados internos com propriedades semânticas que (de alguma forma) representam o mundo externo, o enativismo enfatiza o caráter corporificado e situado da cognição, insistindo que o mental emerge da exploração ativa do organismo em seu meio e que não há um ‘mundo’ a ser representado independentemente do modo como este organismo age e modifica seu meio. Ao contrário, mente e mundo são co-determinados mutuamente a partir do histórico de enação do organismo. Nessa perspectiva, para explicar a ação bem-sucedida (e os erros) de um organismo, basta especificar seus acoplamentos habilidosos em seu nicho, não havendo necessidade de se postular algo como uma representação interna do mundo exterior. Como dizem Varela, Thompson & Rosch no clássico enativista *The Embodied Mind*:

We propose as a name the term *enactive* to emphasize the growing conviction that cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs (Varela; Rosch; Thompson, 1991, p. 9).

Mesmo que o enativismo seja um programa de pesquisa amplo com diferentes vertentes¹, é seguro dizer que elas têm em comum uma certa concepção da mente expressa

¹ Como por exemplo o enativismo autopoietico de Varela, Thompson & Rosch (1991), o enativismo radical de Hutto & Myin (2013) ou o enativismo fenomênico de Noë (2004). Para uma excelente introdução a esse programa de pesquisa em suas diferentes vertentes, ver Rolla (2021).

pelos 4 “E” da cognição – ou seja, que esta é essencialmente corporificada (*embodied*), situada (*embedded*), enativa (*enacted*) e estendida (*extended*). Neste artigo também subscrevemos a essa concepção², algo que nos coloca em um ambiente teórico hostil para qualquer tipo de vocabulário representacional, sob o risco de associação a um ultrapassado “cognitívismo de velha guarda” que Rolla (2021, p. 23-4) coloca da seguinte forma:

Cognitívismo de velha guarda, ou antigo cognitívismo, é a combinação de duas teses: em primeiro lugar, que a mente é essencialmente representacional (representacionalismo) e, em segundo lugar, que processos cognitivos são computacionais (computacionalismo). Segundo essa perspectiva, seria possível entender como a mente funciona completamente à parte das características corporais do organismo e do ambiente em que ele se encontra.

Embora estejamos de acordo com as críticas enativistas ao chamado “cognitívismo de velha guarda”, ainda não é claro para nós que pensar a cognição como corporificada, enativa e situada implique a rejeição de qualquer tipo de vocabulário representacional; ou, dito de outra forma, que introduzir ‘representação’ como um termo teórico útil em alguns contextos nos comprometa com algo remotamente semelhante ao que Rolla (2021) descreve como “cognitívismo de velha guarda”. Mas como nos encontramos em um território hostil a esse tipo de vocabulário, será preciso esclarecer o que entendemos por ‘representação’, porque julgamos que esta noção pode nos trazer alguma vantagem explicativa em certos casos e porque aceitá-la não implica de forma alguma trair esses preceitos metodológicos.

Que fique claro, nosso objetivo neste artigo não é argumentar em favor de algum tipo de representacionalismo. Na verdade, não acreditamos que seja possível sustentar o representacionalismo ou o antirrepresentacionalismo como uma tese generalizada acerca da cognição. Ao contrário, se devemos ou não introduzir uma ‘representação’ como um termo teórico útil é algo que só pode ser estabelecido empiricamente a partir de uma análise caso a caso, considerando as características específicas do organismo, da tarefa e do comportamento a ser explicado.

Nossa estratégia neste artigo consistirá, portanto, em: (i) argumentar que, pelo menos em alguns casos, postular representações pode nos comprar uma dimensão explicativa adicional que não está disponível fora deste vocabulário; e (ii) argumentar que esta noção de representação não contradiz os preceitos metodológicos mais caros ao enativis-

² Embora permaneçamos neutros em relação ao quarto “E”, isto é, a mente estendida, enquanto uma tese ontológica sobre as fronteiras do mental.

mo. Em particular, (a) ela não nos compromete com um mundo pré-dado representado por uma mente pré-dada, nem (b) com um computacionalismo ou representacionalismo generalizados. Isto é, mesmo que a representação seja um recurso teórico útil em certos casos, disso não se segue que ela possa ser aplicada sem restrições a outros casos.

A estratégia de apresentação será a seguinte: na seção (1), delinearemos nossa posição entre o cognitivismo clássico e as abordagens não representacionais como o enativismo. Na seção (2), apresentaremos a distinção entre estados representacionais e estados intencionais (alvos), que será utilizada ao longo de todo o argumento. Na seção (3), apresentaremos a teoria da representação de Robert Cummins, argumentando que ela é capaz de evitar tanto os problemas clássicos quanto os desafios antirrepresentacionais conhecidos. Na seção (4), argumentaremos pela relevância explicativa das representações, sugerindo que elas nos permitem formular teses empíricas que estariam indisponíveis (ou seriam distorcidas) em sua ausência. Na seção (5), antes de proceder para a conclusão, faremos algumas considerações sobre como essas ideias podem ser adotadas em *frameworks* já existentes.

1. Batalha de *frameworks*

Como vimos, é comum caracterizar o cognitivismo clássico pelo seu comprometimento com o computacionalismo e o representacionalismo: a mente cognitiva é constituída por registros representativos do mundo sobre os quais se realizam operações exclusivamente computacionais. Nessa perspectiva, a cognição se esgota no papel de mecanismos computacionais e estados representacionais. Menos comum é enfatizar que o cognitivismo clássico adotava uma concepção específica do que são mecanismos computacionais, e uma concepção igualmente específica do que são estados representacionais. Por exemplo, representações mentais seriam simbólicas e registrariam o mundo num formato proprietário, com forma sentencial (FODOR, 1980)³. A ontologia subjacente a essa forma de registro (objetos, propriedades e relações) é o que tornava plausível descrever processos cognitivos por meio de formalismos lógicos. Vem daí também sua característica mitigação da importância da percepção. Processos perceptuais seriam “anteriores” a esse registro do mundo, e não pareciam ensejar dificuldades de

³ Uma vez que o foco deste trabalho são representações mentais, deixaremos a discussão sobre mecanismos computacionais de lado.

naturalização⁴. A despeito de todos os seus problemas, essas ideias harmonizavam bem com a tese de que a cognição compreende um domínio de pesquisa autônomo e independente do corpo e do ambiente.

A explicitação desse modo específico de compreender o que são representações – e qual seu papel – é fundamental para distinguir bons argumentos contra o cognitivismo clássico de bons argumentos contra representações mentais em geral. Quando a autonomia do domínio cognitivo começou a ser questionada por uma nascente concepção 4E, nem todos se atentaram a isso. Via de regra, adotou-se um apressado pessimismo quanto à possibilidade de convivência pacífica entre estados representacionais e os preceitos da cognição 4E. Isso se mostra ainda hoje na resistência de alguns enativistas em aceitar qualquer concepção de representação mental. O temor parece ser o de que, ao permitir o uso de recursos explicativos de berço cognitivista, abra-se espaço para a adesão sub-reptícia de princípios metodológicos indesejados.

Um exemplo concreto de como esse receio se manifesta está na relação entre cognição, ação e percepção. A cognição 4E entende que a relação organismo-ambiente deve ser tomada como unidade de análise elementar. Resulta daí um abandono do que Hurley (2001) caracterizou como modelo “sanduíche” da relação entre percepção, mente e ação. A mente cognitiva não seria um espelho do mundo posicionado entre *inputs* perceptuais e *outputs* comportamentais, mas algo que emerge da interação contínua e direta entre ação e percepção. Se reintroduzirmos alguma concepção de representação, não poderia isso nos forçar a uma retomada do modelo sanduíche?

Trabalhos como o de Piccinini (2021) e Clark (2016) respondem a essas preocupações acomodando preceitos caros à cognição 4E sem abrir mão de representações. Porém, isso tem se mostrado insuficiente. Parte do problema é que persiste o clima de embate entre *frameworks* que almejam autossuficiência. Isso gera uma luta por espaço que só acabará se e quando uma engolir a outra. Esse embate é alimentado por diagnósticos equivocados acerca do que representações são e de como são usadas. A preocupação com um possível retorno ao modelo sanduíche, por exemplo, não diz respeito à existência de conteúdo representacional. Ela concerne se, quando e como esse conteúdo é explorado pelo sistema. No decorrer do artigo, buscaremos mostrar como essa distinção pode ser feita. Por ora, importa notar que, se estivermos no caminho certo, talvez seja possível mudar a relação entre os *frameworks* atuais: em vez de inflamar a disputa por espaço, será possível ampliar o conjunto de preceitos partilhados.

4 Até onde se sabe, capacidades perceptuais nunca motivaram nenhum tipo de dualismo ao longo da história.

Mas por que proponentes do enativismo, da psicologia ecológica e/ou de quaisquer abordagens não representacionais deveriam nos dar ouvidos? Há muito trabalho sendo feito para mostrar como capacidades que parecem demandar representações mentais podem ser explicadas mesmo na sua ausência. Certamente nenhum desses trabalhos é conclusivo, mas por que abrir mão do norte que perseguem? Não é isso que queremos sugerir. O ponto é mostrar que nenhum preceito essencial à cognição 4E precisa ser abandonado, mesmo que ao fim do dia as representações se mostrem indispensáveis em alguns casos. Se formos bem sucedidos, portanto, teremos em mão uma proposta legitimamente conciliatória e (esperamos) bastante atraente.

2. Preliminares: a distinção entre representações e alvos

Há algum tempo, Cummins (1996) percebeu que os representacionistas tendem a ignorar uma distinção entre duas dimensões presentes nos mecanismos cognitivos que exploram representações. Quando um sistema utiliza uma representação para lidar com algum elemento do mundo, essa aplicação pode ser analisada tanto pela via semântica quanto pela via funcional. De um lado, a dimensão semântica responde por aquilo que um dado estado ou processo efetivamente representa. Do outro, temos a dimensão do alvo a que a representação é aplicada, ou seja, aquilo que o sistema pretende explorar pelo uso da representação como um modelo do alvo. A noção de alvo é funcional: ela especifica aquilo que um dado mecanismo ou processo tem a função de representar em uma dada circunstância.

Essa distinção é facilmente identificável no âmbito agencial. Suponha-se um grupo de amigos a viajar por um país desconhecido nos anos 1990. Igor recebe a tarefa de providenciar um mapa do bairro em que se encontram para que todos possam localizar um museu que desejam visitar. Ele não encontra nenhum mapa impresso à venda, mas coleta informações diversas dos habitantes locais. Com base nelas, desenha a estrutura das ruas, suas intersecções, acrescenta alguns pontos de referência e, por fim, especifica a localização do museu. Ele reencontra os amigos e lhes apresenta o mapa, que será então utilizado por todos para chegar ao local desejado. Esse exemplo permite ver com clareza que, qualquer que seja o conteúdo do mapa, ele será tomado pelos amigos como um *mapa do bairro da cidade em que se encontram*. Esse era o alvo de Igor em virtude da função que lhe foi atribuída.

Contudo, o mapa produzido pode não representar adequadamente a estrutura do bairro da cidade. Um dos habitantes consultados pode ter mentido, ou talvez se confun-

dido, ainda que de boa fé. Seja como for, todas as inferências feitas pelo grupo partirão do princípio de que o mapa representa o bairro da cidade em que eles se encontram. Essa é a distinção usada por Cummins (1996) no âmbito sub-pessoal: um mecanismo pode ter como função representar um dado alvo X. A representação que ele produz será tomada pelos mecanismos consumidores como sendo de X, mesmo que ele produza algo que não represente X acuradamente.

Neste cenário, estados ou processos representacionais explorados por sistemas cognitivos se relacionam com o mundo por dois caminhos: primeiro, por terem conteúdo (i.e. representarem algo); segundo, por serem explorados como modelos de algo. Essa distinção permite que cada dimensão seja objeto de uma teoria específica. Uma teoria do conteúdo representacional trata da dimensão semântica, e uma segunda teoria é responsável por tratar do modo como um dado organismo fixa os alvos aos quais aplicará as representações produzidas. Para Cummins (1996), essas teorias podem e devem ser independentes. Perguntar pela cidade que um mapa representa é diferente de perguntar pela cidade que se está tentando explorar por meio daquele mapa. Da mesma forma, o modo como o conteúdo de um estado ou processo é fixado independe do modo como o organismo fixa seus alvos.

Com efeito, o caráter intencional de um determinado estado ou processo não é fruto do conteúdo representacional explorado, mas sim do alvo fixado. A intencionalidade é um fenômeno mais amplo que a intencionalidade representacional. Depreende-se disso que, em larga medida, o debate contemporâneo entre abordagens representacionistas e não representacionistas pode ser compreendido como um debate sobre o modo como um dado organismo consegue fixar em certos alvos. Um alvo pode ser fixado tanto pela especificação funcional de um mecanismo inato, quanto pela especificação de um processo que emerge num arranjo muito particular envolvendo o aparato cognitivo, o corpo e o ambiente do organismo. Por exemplo, a tese de que organismos exploram informação ecológica é uma tese sobre fixação de alvos. Ela nos diz que sistemas (organismos) podem fixar como alvos certas estruturas complexas presentes no ambiente, inclusive estruturas dinâmicas, como as que se articulam ao longo do tempo. Evidentemente, cognitivistas e enativistas podem divergir sobre quais os alvos utilizados por um dado organismo e como ele os fixa. Mas essa divergência não necessariamente diz respeito à aplicação de representações no modo como esses alvos são explorados. Representar não é um modo distinto de fixar alvos. Representações não são um meio de caracterizar a sensibilidade a características do ambiente, mas sim uma estratégia à disposição do organismo para explorá-las e armazenar informação sobre elas.

Sob a luz dessa distinção, podemos nos voltar agora para as razões que tipicamente justificam o pessimismo antirrepresentacionista. Costuma-se explorar dois flancos: primeiro, uma aparente dificuldade de naturalização. Segundo, o papel aparentemente trivial da representação na explicação científica, geralmente em virtude da suposta inércia causal do conteúdo semântico. Ambas conduzem a uma negação da realidade do conteúdo representacional ou a abordagens deflacionadas, como a de Egan (2020). Mas elas podem ser contornadas, e é disso que nos ocuparemos agora.

3. Representações: o que são?

Nos últimos anos, a dificuldade de naturalizar estados representacionais se condensou na formulação de Hutto & Myin (2013), conhecida como *o problema duro do conteúdo*. Uma síntese do argumento pode ser encontrada em Rolla (2023, p. 213):

[...] de acordo com cognitivistas, representações mentais são portadoras de informação semanticamente carregada [...]. Porém, o único tipo de informação encontrada na natureza é a covariação. [...] Não podemos inferir que um dos termos em uma relação de covariação representa o outro. Números de anéis no tronco da árvore não representam a sua idade — a representação aqui é imputada por nós [...]. Por si só, estados naturais são piamente quietistas e não dizem nada sobre ninguém. [...] [P]or que seriam os estados cerebrais diferentes dos demais estados naturais? [...] Ou o cognitivista aceita que existe cognição sem conteúdo (acarretando o enativismo) ou nos passa um cheque sem fundo com a promessa de que um dia a física do futuro vai descobrir conteúdos semânticos em estados naturais. A segunda via é naturalisticamente temerosa. Donde se segue que o cognitivista deve dar o braço a torcer [...] (Rolla, 2023, p. 213)⁵.

O alvo do argumento são teses ancoradas na covariação confiável entre estados e propriedades: se *F* covaria com *G*, *F* *significa* *G*. Esse é o tipo de informação a que Dretske (1981; 1986) e outros recorrem para ancorar conteúdo representacional. Mesmo antes da formulação de Hutto & Myin (2013), os desafios dessa abordagem eram bem conhecidos⁶. Porém, não é necessário aprofundarmos a discussão. Independentemente do que se conclua

⁵ Embora esse ponto seja tangencial à discussão, é importante salientar que mesmo o eventual sucesso do argumento não acarretaria o enativismo como se afirma. A situação resultante seria compatível com toda sorte de teorias não representacionais, a exemplo do conexionismo eliminativista ou da “teoria sintática da mente” de Stich (1983), que questiona a relevância explicativa do conteúdo representacional sem abrir mão do modelo sanduíche da ação e percepção ou de modelos computacionais clássicos.

⁶ Por exemplo, Cummins (1996), Perlman (2002) e Ramsey (2007).

acerca de teorias baseadas em covariação, o argumento não cumpre o prometido porque é falso que a covariação seja a única relação natural explorável. Há também a relação de isomorfismo estrutural (ou homomorfismo, entendido aqui como um isomorfismo que acomoda gradações). Para Cummins (1996) e Swoyer (1991), a dimensão semântica dos estados ou processos representacionais pode ser completamente acomodada nessa relação: representação nada mais é do que a relação matemática de isomorfismo. Isso vale para qualquer estado ou processo físico, de mapas impressos a estados cerebrais. Uma estrutura A representa uma estrutura B na medida em que A é isomórfica a B. Com efeito, elementos de A representam elementos de B, relações entre esses elementos presentes em A representam relações entre esses elementos em B, e assim por diante. Temos assim uma semântica para *representações estruturais*.

A tese é simples, mas faz emergir dúvidas sobre como esse conteúdo pode cumprir o papel que se espera dele. Uma primeira preocupação diz respeito a uma aparentemente excessiva liberalidade: há estruturas isomórficas por toda parte. Isso significa que A pode representar não apenas B, mas também C, D, E, e assim por diante, indefinidamente. Um mapa contendo a estrutura das ruas de um dado bairro de Belo Horizonte (BH) representa aquelas ruas de BH, mas pode também representar as ruas de outras cidades, bem como qualquer outra estrutura existente que calhe de ser isomórfica. Se esse é o caso, o conteúdo representacional será sempre não único. Porém, é aqui que a distinção entre alvo e conteúdo começa a render dividendos. A pressão pela determinação do conteúdo se faz sentir tão somente no âmbito da fixação de alvos. Mais precisamente, o que pode fazer parecer com que a liberalidade da relação de isomorfismo seja inadequada é a confusão entre dois sentidos diversos em que podemos compreender a afirmação “A representa B”:

- (1) a estrutura A é isomórfica à estrutura B;
- (2) a estrutura A é explorada por um sistema como um modelo de B.

Cummins (1996) se distingue ao usar o termo “representação” no sentido (1), pois o significado usualmente aplicado na literatura é (2). Considere, por exemplo, a seguinte caracterização de Hutto & Myin (2013, p. 62):

To qualify as representational, an inner state must play a special kind of role in a larger cognitive economy. Crudely, it must, so to speak, have the function of saying or indicating that things stand thus and so, and to be consumed by other systems because it says or indicates in that way.

Tal caracterização claramente remete a (2). Ocorre que (2) confunde a dimensão semântica e a dimensão funcional associada à fixação de alvos. Ela não nos dá aquilo em virtude de que uma dada estrutura pode ser considerada representacional (para isso, o isomorfismo é suficiente). O que ela expressa são as condições a caracterizar mecanismos que exploram um dado estado representacional. Em outras palavras, ela caracteriza casos de *aplicação* de representações a alvos funcionalmente determinados.

Se prescindirmos da distinção entre alvo e representação, contudo, seremos tentados a concluir que só poder haver conteúdo onde há aplicação desse conteúdo a um alvo determinado. Isso é um problema porque, numa economia representacional, estas dimensões exercem papéis explicativos distintos. A dimensão semântica não se ocupa de identificar o que o mecanismo está tentando fazer, mas, sim, de contribuir para uma compreensão da performance resultante: por que o organismo foi bem (ou mal) sucedido? Se um mecanismo tem a função de fornecer o mapa de um bairro das ruas de BH, então seus consumidores irão tomar as representações fornecidas como sendo de BH. Ainda que ela calhe de ser isomórfica a um sem número de outras estruturas, isso não afeta a explicação da performance obtida. Afinal, são as ruas de BH que constituem a norma contra a qual a acurácia da representação produzida deve ser mensurada.

Nessa abordagem, nenhum dos problemas tradicionais relacionados à naturalização do conteúdo precisa nos preocupar. Não é preciso, nem faria sentido, reduzir isomorfismo (ou homomorfismo) a qualquer outra concepção de informação, visto ser ela uma relação perfeitamente natural. Não há, portanto, mistério sobre como uma representação se ancora no mundo (*grounding problem*). Não há problema de coordenação entre conteúdo e veículo, e tampouco há dúvidas sobre se e como eles podem ter eficiência causal: representações guiam o processamento em virtude de seu conteúdo, e seu conteúdo é dado pela sua estrutura. Uma vez que tanto os poderes causais quanto o conteúdo são dados pela estrutura, não há necessidade de “interpretadores” desse conteúdo. Consumidores das representações são causalmente sensíveis às estruturas, da mesma forma que um cadeado é causalmente sensível a uma chave. Além disso, não há real problema relacionado à determinação do conteúdo. Como vimos, o caráter não único do conteúdo estrutural não o impede de cumprir adequadamente o papel que o cognitivismo lhe prescreve. Afinal, a pressão por determinação se dá sobre a dimensão da fixação de alvos.

No cenário articulado, o problema de como identificar mecanismos que exploram representações (i.e. que exploram propriedades de uma estrutura a fim de lidar com um alvo do mundo), emerge como empírico. Deve-se determinar caso a caso se, e quando, um dado mecanismo explora estados representacionais no exercício de sua função. Conside-

re, por exemplo, o contraste com mecanismos não representacionais que exploram sinais emitidos por detectores. Tais sinais são indicadores da presença de alguma propriedade ou estado de coisas no mundo. Exemplos clássicos são sinais de detectores de temperatura em sistemas de ar condicionado, de nível de combustível ou pressão de óleo em veículos, e de certos padrões no campo visual, a exemplo das células no córtex visual primário. Esses sinais também têm alvos, mas não os representam. Seu conteúdo é tão somente uma indicação de presença que advém da especificação funcional do mecanismo emissor. Sabemos estar com pouco combustível diante de um sinal emitido por um mecanismo com essa função. Dissociar o sinal daquele detector faz com que a informação seja perdida. Indicadores e representações são, nesse sentido, fontes distintas de tipos distintos de conteúdo.

Como podemos distinguir casos em que um organismo está efetivamente explorando uma representação dos casos em que ele está explorando outros tipos de recurso, tais como sinais de indicadores? A mera detecção de estruturas instanciadas no aparato cognitivo é insuficiente. Elas podem caracterizar o que Cummins denominou “representações não exploradas” (Cummins *et al.*, 2010), isto é, estruturas isomórficas a algo, mas que não são explorados em virtude desse isomorfismo⁷. A distinção é importante porque, como nos lembra Facchin (2021), mesmo mecanismos indicadores podem apresentar estados isomórficos ao que detectam no mundo⁸. Mecanismos que consomem sinais de indicadores, contudo, não são sensíveis à essas estruturas. Portanto, é preciso determinar se a capacidade ou performance estudada envolve processos causalmente sensíveis às estruturas instanciadas. Shea (2018) sugere que o foco desse tipo de investigação deve estar na correlação entre o grau de acurácia representacional e o comportamento resultante. Na medida em que representações estruturais mais acuradas resultam em comportamentos mais adequados, é plausível inferir que a estrutura exerce papel relevante.

7 A possibilidade de conteúdo representacional não explorado é usada por Cummins (1996; 2000) para criticar abordagens teleosemânticas como a de Millikan (1987). A tese central é a de que a teleosemântica inverteria a ordem explicativa, fazendo com que a atribuição de conteúdo derive da adaptatividade, quando o que queremos é explicar como foi possível que a exploração de um dado conteúdo tenha se mostrado adaptativa. Curiosamente, Hutto & Myin (2013) citam Cummins *et al.* (2010) explicitamente ao fazer suas próprias críticas a Millikan, sem, contudo, abordar a solução oferecida por Cummins no mesmo artigo (a mesma que aqui desenvolvemos).

8 Isso leva Facchin (2021) a concluir equivocadamente que essas estruturas não representam. Elas representam, sim, aquilo a que são isomórficas. O que está em jogo é a existência de algum mecanismo sensível a essas estruturas, isto é, consumidores.

Embora seja um passo na direção certa, Shea (2018) parte de uma suposição problemática: a de que maior acurácia representacional implica melhor performance. Mas sucesso semântico não implica sucesso comportamental ou funcional. No caso de criaturas limitadas em recursos (como nós), representações pouco precisas podem ser a razão pela qual um dado comportamento resultante se mostrou adaptativo ou apto o suficiente para ser adotado de forma robusta. Acurácia representacional costuma exigir maior dispêndio de recursos do organismo na tentativa de explorá-la, e isso pode resultar em comportamento inadequado. Além disso, uma resposta correta à pergunta “qual o grau adequado de acurácia?” é claramente dependente das circunstâncias. O que seria considerado um erro grosseiro numa situação pode ser suficiente ou vantajoso em outra. Situações que acomodem tempo e memória suficientes podem se beneficiar de uma precisão maior, e situações que impõem constrangimentos de tempo e memória (o organismo está fugindo de um predador) podem mitigar o grau de acurácia considerado ideal. Pela mesma razão, mecanismos mentais podem ser considerados funcionais ou eficientes, ainda que envolvam frequentes erros semânticos (i.e. imprecisões). Com efeito, os critérios adotados para identificar casos em que um organismo explora representações não podem pressupor correlação entre sucesso funcional, comportamental e semântico.

Embora isso complique as coisas, obtemos em troca um grau de liberdade crucial. Modelos que busquem explicar como funcionam os mecanismos subjacentes a uma dada capacidade cognitiva podem articular essas diferentes dimensões. Pode-se formular diferentes hipóteses empíricas sobre o alvo adotado, o conteúdo produzido (se algum), como esse conteúdo é explorado (se o for), e qual o comportamento resultante em diferentes circunstâncias. Ainda que diferentes modelos possam explicar satisfatoriamente o mesmo tipo de comportamento, a articulação dessas dimensões amplia o conjunto de possíveis efeitos incidentais associados a cada modelo. Por “efeito incidental”, queremos indicar uma espécie de efeito colateral do modelo. Trata-se de um efeito causal que não participa da explicação da capacidade que se busca modelar, mas que pode ser útil para decidir qual tese evoca o modelo mais adequado dentre os candidatos disponíveis (Cummins, 2010). Para deixar esse ponto mais claro, vamos tratar primeiro de um exemplo hipotético e depois de um exemplo mais concreto.

Suponha-se que queiramos explicar como uma calculadora é capaz de multiplicar números e tenhamos dois modelos concorrentes. Um deles supõe que a calculadora realiza somas sucessivas: o resultado de 25×15 é dado por 15 somas sucessivas ($25+25+25\dots$). O outro modelo supõe que a calculadora faça uso do algoritmo de produtos parciais que todos aprendemos na escola. Ainda que ambos predigam o mesmo *output*, o primeiro

modelo o faz por meio de uma composição simples, enquanto o segundo explora uma relação estrutural existente no esquema arábico (note a importância da posição em que os produtos parciais são armazenados, e também como esse algoritmo está indisponível para a notação romana). Isso gera efeitos incidentais distintos. O primeiro modelo é mais sensível ao tamanho do número (15, 30, 45...), resultando numa determinada trajetória a descrever o consumo de tempo e memória. Já o segundo modelo prediz uma trajetória de consumo de recursos que, embora também sensível ao tamanho dos números envolvidos, se mantém estável quando o número de dígitos é o mesmo, apresentando saltos somente quando há um incremento destes (15, 115, 1015...). Temos assim um efeito empírico que pode ser mensurado para decidir qual dos modelos descreve melhor a capacidade que se buscava explicar, mesmo em casos onde o comportamento resultante não varie.

Um exemplo do mundo real pode demonstrar a utilidade de efeitos incidentais de modo ainda mais claro. Newen & Vosgerau (2020) apontam para um caso interessante envolvendo ratos e sua capacidade de articular informações sobre a organização espacial de labirintos, tipos de comida e períodos do dia. Eles demonstram capacidade de compreender, por exemplo, que quando um certo tipo de comida (salgada) foi disponibilizada em certa posição do labirinto na parte da manhã, haverá um outro tipo de comida (doce) em um outro dado local do labirinto na parte da tarde (Crystal, 2013; Panoz-Brown *et al.*, 2016). Os autores defendem que a melhor explicação disponível envolve articulações entre representações de tipos de comida, de períodos do dia e, claro, de localização espacial. Mais do que a verdade ou falsidade dessa explicação, o que realmente nos importa notar aqui é a argumentação utilizada pelos autores contra hipóteses alternativas:

The informational state of rats which have learned to behave according to a conditional in the maze is best characterized as structured into components of <object-type; location; time>. The alternative would be to presuppose a high number of independent, non-structured dispositions which need to include all the possible permutations of associations between a starting state of affairs and a type of behavior. And these dispositions would need to be learned independently of each other, since there would be no common component to be taken over (Newen; Vosgerau, 2020, p. 184).

Essa argumentação depende crucialmente da articulação de um efeito incidental à capacidade sendo explicada. Embora seja concebível que o tipo de comportamento exibido possa ser explicado por um conjunto razoavelmente grande de diferentes *affordances*, essa alternativa requer pelo menos uma *affordance* independente para cada

permutação da estrutura representacional. Não sendo esse o caso, o comportamento do rato permaneceria parcialmente inexplicado. O problema com esse requisito é que ele implica a manifestação de um efeito incidental incompatível com as atuais evidências empíricas (CRYSTAL, 2013; PANOZ-BROWN *et al.*, 2016). Ratos apresentam uma curva de aprendizado rápida e com razoável flexibilidade, e isso é observado especialmente em casos de variantes conservadoras das situações já conhecidas. Eles podem, por exemplo, compreender que um labirinto foi parcialmente rearticulado ou que, em certas condições, o conteúdo de uma parte do labirinto pode ser considerado irrelevante para determinar a provável presença de alimento num certo período do dia. Um conjunto não estruturado de *affordances* ou disposições implicaria uma trajetória de aprendizado mais lenta e, possivelmente, mais problemática. Em particular, ela seria sujeita a falhas comportamentais distintas das observadas sob a hipótese representacional, especialmente no caso de permutações mais complexas entre diferentes elementos da tríade objeto-local-tempo.

Temos então um cenário em que a pergunta “o que são representações?” é claramente distinta da pergunta “como elas são exploradas?”. Representações são isomorfismos. Representações mentais são isomorfismos que ocorrem no interior do aparato cognitivo. A pergunta sobre como elas são exploradas é, portanto, uma pergunta sobre como determinados sistemas fazem uso delas para lidar com os alvos a que têm sensibilidade. Segue-se que não é preciso formular um conjunto de condições necessárias e suficientes para estabelecer quais mecanismos podem ou não ser considerados representacionais. O que se deve fazer é buscar os casos em que modelos envolvendo representações caracterizam a melhor explicação disponível face às evidências. E o que caracteriza a melhor explicação no caso de um mecanismo pode ser diferente em outro. Como se viu, ainda que possam haver casos em que dois ou mais modelos sejam capazes de explicar uma dada capacidade, eles podem fazê-lo por caminhos distintos, e efeitos incidentais podem ser utilizados para identificar qual hipótese é melhor acomodada pelos dados.

4. Representações: modo de usar

O próximo passo é argumentar que a atribuição de conteúdo às estruturas exploradas é capaz de cumprir um papel explicativo não trivial. Precisamos mostrar que representações fazem diferença na explicação de capacidades cognitivas e que, portanto, não estão apenas nos olhos de quem vê. Representações são reais e podem ter papel explicativo relevante porque permitem a formulação de teses empíricas indisponíveis na sua ausência. Essa dimensão explicativa adicional se caracteriza pela distinção entre erro

representacional e erro de processamento/exploração. Considere novamente um exemplo de como representações não mentais podem ser exploradas em nível agencial:

- (a) Igor errou o caminho porque o mapa era excessivamente impreciso; e
- (b) Igor errou o caminho porque usou o mapa de cabeça pra baixo.

A tese (a) caracteriza inacurácia no modo como a informação foi representada. É um exemplo de erro representacional. Já a tese (b) caracteriza um erro no modo como a informação foi explorada. Falhas mecânicas, erros de engajamento ou acoplamento, bem como condições aquém das ideais, são exemplos típicos de erros desse tipo. Erros representacionais independem da aptidão com que a informação é explorada ou de problemas no veículo (quando um mapa físico é danificado, por exemplo). O que os torna possíveis é a relação *direta* entre a estrutura a que a representação é isomórfica (conteúdo) e a estrutura a que ela é aplicada (alvo). Assim, eles desdobram toda uma classe de hipóteses empíricas: as que envolvem a articulação da diferença entre conteúdo e alvo. A tese de que Igor errou o caminho por imprecisão do mapa é distinta da tese de que ele o explorou mal. Ainda que resultem em comportamentos idênticos, elas têm efeitos incidentais distintos.

Outro sinal de que erros representacionais desdobram uma classe particular de hipóteses empíricas está na sua independência. Por si mesma, a imprecisão semântica não implica nem sugere insucesso funcional ou comportamental. Por isso, no exemplo (a), é necessário caracterizar o erro comportamental de Igor (errar o caminho) como fruto de imprecisão excessiva, isto é, de um grau de inacurácia prejudicial aos seus objetivos. Contudo, abrir mão de acurácia em virtude de tratabilidade é frequentemente vantajoso. Quando desenhamos um mapa simples, ou usamos um modelo simplificado de um domínio complexo para compreendê-lo melhor, buscamos nos poupar do esforço de lidar com grande quantidade de detalhes que ampliam a acurácia, mas que são irrelevantes para o resultado. A mesma estratégia está disponível para a natureza. Com recursos limitados, uma representação simplificada da dinâmica dos movimentos do predador num mecanismo sub-pessoal pode ser a chave para que a presa tenha tempo de escapar.

Para um exemplo concreto, podemos retomar os experimentos anteriormente citados envolvendo ratos em labirintos. Newen & Vosgerau (2020) argumentam que o modo flexível e rápido como os ratos aprendem é melhor explicado representacionalmente. Ou seja, a trajetória desse aprendizado é melhor caracterizada nos termos da correção da imprecisão representacional (i.e. redução da diferença entre conteúdo e alvo). O ponto,

contudo, não é o de que representações mentais sempre caracterizarão a melhor explicação. Trata-se de ressaltar que, na ausência de boas razões *a priori* para descartá-las, a questão é empírica. Determinar se o exercício de uma capacidade cognitiva envolve erros do tipo (a), do tipo (b) ou de alguma articulação mais complexa entre eles, requer avaliação caso a caso. Assim, o problema de *frameworks* que rejeitam representações mentais é que descartam ou distorcem uma classe de hipóteses empíricas por meios não empíricos. A empreitada científica acaba indevidamente constrangida, fazendo parecer que a única caracterização empiricamente plausível do comportamento é a de que se cometeu um erro ao explorar a informação disponível, i.e., um erro do tipo (b).

Contudo, antirrepresentacionistas não são os únicos tentados a dizer que o que parecem erros do tipo (a) são, no fundo, erros do tipo (b). Vários autores representacionistas acreditam (erroneamente) que esse é o único caminho plausível para uma teoria naturalista da representação. Entender onde essa tentativa começa é importante para apreciar a robustez que a distinção entre conteúdo e alvo nos compra. Considere o que John Haugeland (1998, p. 309–10) nos diz acerca do alcance explicativo de normas biologicamente sedimentadas:

[...] there is another important distinction that biological norms do not enable. That is the distinction between functioning properly (under the proper conditions) as an information carrier and getting things right (objective correctness or truth), or, equivalently, between malfunctioning and getting things wrong (mistaking them). Since there is no other determinant or constraint on the information carried than whatever properly functioning carriers carry, when there is no malfunction, it's as "right" as it can be. In other words, there can be no biological basis for understanding a system as functioning properly, but nevertheless misinforming [...].

Haugeland (1998) argumenta que normas biológicas não permitem falar de organismos que “erram”, no sentido de que estão operando adequadamente, em condições ideais e, mesmo assim, operam com informações inaccuradas. Ele demonstra com convicção de que erros representacionais só podem ser naturalizados por meio da redução a erros de outra ordem.⁹ Novamente, a análise de Cummins (1996) nos permite enxergar

⁹ O texto de Haugeland (1998) foi escrito em um momento ainda dominado pelo adaptacionismo ancorado na síntese evolutiva moderna em biologia. A síntese evolutiva estendida fornece novas dimensões nas quais talvez possamos ancorar erros perceptuais. Para um exemplo, vide Carvalho & Rolla (2020). Contudo, importa notar que esse tipo de erro é diferente do que se busca articular aqui: ele não é semântico e não nos compra uma dimensão representacional não trivial. Nesse sentido, ele se insere no mesmo grupo das abordagens aqui tratadas, que buscam fundamentar o erro semântico em elementos não semânticos. Além

as razões dessa dificuldade com clareza. Começemos com a hipótese mais simples, em que o conteúdo de uma representação é dado pelo alvo a que ela é aplicada. Esse é o caso das semânticas de papel conceitual ou inferencial, quando aplicadas à determinação de conteúdo representacional. Nelas, o conteúdo atribuído a uma certa representação é fruto direto do seu padrão de uso. Tais padrões determinam os alvos a que uma representação é aplicada pelo sistema e são determinados pelas relações mantidas com as demais representações que o sistema é capaz de instanciar. Assim, tanto o alvo quanto o conteúdo de um estado representacional são dados pelo modo como ele é usado. O conteúdo é, por definição, idêntico ao seu alvo.¹⁰ Isso impossibilita o erro representacional, que se caracteriza justamente pela distinção entre a estrutura de um alvo e a estrutura utilizada como modelo daquele alvo.

Para escapar dessa limitação, é preciso criar um vão entre representação e alvo. Isso é exatamente o que a distinção de Cummins (1996) nos compra. Mas se a ignorarmos, como fazem os teóricos mais conhecidos, seremos forçados a ancorar esse vão em algum elemento não semântico. Como Perlman (2002) sintetizou, esses elementos adicionais buscam dividir as aplicações de representações em duas classes distintas: as que fixam conteúdo, e as que aplicam conteúdo previamente fixado, sendo as últimas sujeitas a erro. Como exemplo, considere a abordagem teleológica de Millikan (1987). Ela busca acomodar a dimensão do erro representacional na adaptatividade. O que distingue aplicações corretas e incorretas é o papel adaptativo do conteúdo na história evolutiva da espécie. Segue-se que toda representação responsável por estabelecer comportamento que se mostrou adaptativo é considerada correta. Em outras palavras, a adaptatividade é responsável por distinguir os casos de uso que fixam conteúdo dos casos de uso que aplicam o conteúdo fixado, e são os últimos que acomodam a possibilidade de erro. Mantém-se, portanto, a doutrina de que o conteúdo representacional de um determinado estado é fruto do uso que é feito dele no interior do sistema.

O resultado não é um erro representacional, mas a tentativa – análoga à do antirrepresentacionista – de reduzi-lo a outro tipo de erro. Millikan (1987), por exemplo, fundamenta a correção semântica na correção comportamental (casos que geraram comportamento adaptativo). Mas como vimos na discussão anterior, sucesso semântico, funcional e comportamental são independentes. Se aceitarmos que a dimensão semântica

disso, vale notar que o cognitivismo é compatível com a síntese evolutiva estendida, conforme demonstra o projeto de Heyes (2018).

¹⁰ Note que essas considerações dizem respeito ao uso de semânticas de papel conceitual como teorias do conteúdo representacional. Não está em discussão o uso desse tipo de teoria na determinação de atitudes.

é parasitária das demais, então seremos forçados a concordar com Haugeland (1998): a biologia não dá espaço para mecanismos que estejam exibindo comportamento apto e que não obstante operem com informações inaccuradas. Afinal, só haverá erro semântico se houver também erro funcional ou comportamental. Esse é um problema grave, pois o objetivo do representacionista é explicar performances funcionais e comportamentais com a ajuda (ainda que não exclusiva) de conteúdo representacional. Mas ele acaba invertendo a ordem e explicando a atribuição de conteúdo a partir das performances funcionais e comportamentais.¹¹ Não admira o antirrepresentacionista sentir-se à vontade para perguntar: se o erro funcional e comportamental já estão explicados, por que acrescentar representações a essa história? Em síntese: é a ausência de uma noção de erro puramente representacional que abre o flanco para que o papel explicativo da dimensão representacional seja trivializado.

Para resistir ao desafio antirrepresentacionista, portanto, é preciso mostrar como a dimensão representacional pode ser estabelecida sem parasitar as demais. Felizmente, já temos tudo o que precisamos. Dado o modo independente como conteúdo representacional e alvo são determinados, segue-se que eles podem divergir. A pressão que nos faz tentar ancorar erro semântico em erros funcionais ou comportamentais sequer emerge, pois não é preciso fabricar um vão entre usos que fixam e usos que apenas exploram conteúdos representacionais. Assim, o vão entre alvo e conteúdo consegue cumprir a demanda de acomodar uma dimensão explicativa (i.e., um espaço para formulação de teses empíricas), indisponível para abordagens não representacionais. *Pace* Haugeland (1998), normas biológicas podem, sim, ancorar o tipo de erro em que um mecanismo (ou um agente) pode atuar em condições ideais, sem problemas de funcionamento e, mesmo assim, estar na posse de informações erradas.

Nesse cenário, ainda é possível que todas as teses empíricas envolvendo representações sejam falsas. Porém, superados os problemas de naturalização, não se pode mais rejeitá-las por princípio, pois sua presença e efeitos são agora objeto de investigação empírica. Nos casos em que elas participarem da melhor explicação disponível face às evidências, não haverá razão para prescindir delas. Por isso é importante que a *framework* adotada permita a articulação e a posterior verificação de hipóteses dessa natureza.¹² Seria

11 Note como a crítica é também aplicável a outras abordagens, como à de Dretske (1981; 1986)

12 Convém notar que este é um caso em que o princípio da parcimônia pode levar a engano. A melhor explicação para uma capacidade simples tomada isoladamente pode levar a problemas quando ela é articulada junto a capacidades mais complexas ou demandantes de recursos explicativos adicionais. É preciso encontrar o modo mais parcimonioso de acomodar a participação de certa capacidade em todo o conjunto

inadequado permitir que todo um conjunto de possibilidades empíricas seja descartado apenas em função da preferência por uma abordagem ou outra. É melhor que a discussão se dê caso a caso, mecanismo a mecanismo.

5. Revoluções incompletas

Nas ciências cognitivas, a ausência de unidade sempre foi a regra. Não apenas existem *frameworks* diversas, como cada uma delas pode operar com diferentes articulações de arquiteturas, modelos e princípios. Não falta quem busque formular princípios unificadores, mas o risco de distorção das evidências disponíveis não deve ser minimizado. A insistência em boas práticas científicas que persigam diferentes estratégias parece ser o melhor remédio. Foi assim que aprendemos, paulatinamente, que o campo das ciências cognitivas é fértil em revoluções incompletas, isto é, ideias que prometiam uma reviravolta radical no modo como compreendemos a mente, mas que sob escrutínio revelaram-se tão somente reformistas.

Considere, por exemplo, os desdobramentos do processamento preditivo (CLARK, 2016). Essa abordagem concebe o aparato cognitivo como sendo essencialmente bayesiano e advoga extenso uso de modelos neurais generativos (HINTON, 2007) como a ferramenta ideal para modelar essas capacidades. Tais modelos são frequentemente anunciados como capazes de contar uma história radicalmente nova sobre a cognição. História essa que promete unificar as explicações dos mecanismos envolvidos na ação, percepção e raciocínio. Contudo, modelos generativos compreendem apenas uma das formas possíveis de se modelar sistemas preditivos. Em um trabalho provocativo, Cao (2020) argumenta que modelos generativos e não generativos são matematicamente equivalentes (ao menos em certos casos centrais, como o da percepção), e isso a ponto de não haver efeitos incidentais claros que permitam decidir entre as duas alternativas. Além disso, há também argumentos plausíveis sugerindo que o tipo de inferência sobre o qual sistemas preditivos repousam podem ser modelados em arquiteturas clássicas (PIANTADOSI; JACOBS, 2016). Não bastasse, há também quem aponte equivalências entre os resultados obtidos pela explicação de certas capacidades a partir de modelos generativos e modelos dinâmicos (ANDERSEN; MILLER; VERVAEKE, 2022).

de capacidades cognitivas. Isso pode levar à escolha de teses que demandem mais recursos, mesmo no caso de capacidades relativamente simples, a exemplo do que ocorre no exemplo dos ratos.

Com efeito, temos um cenário em que se torna paulatinamente mais difícil distinguir o que há de realmente inovador no processamento preditivo, ou o que exatamente a adoção de modelos generativos nos compra que já não pudéssemos comprar com outras arquiteturas. Não se busca aqui sugerir que a resposta a essas perguntas é negativa ou que o processamento preditivo caminha inevitavelmente rumo a uma dissolução em abordagens já existentes. Antes, o objetivo é ilustrar o tipo de escrutínio e dialética a que boas ideias costumam ser submetidas no âmbito das ciências cognitivas.

Assim, ainda que a revolução prometida se mostre incompleta, esse processo pode resultar em alterações relevantes no modo como as ciências envolvidas são praticadas, e estas tendem a ser assimiladas. O resultado são *frameworks* mais ricas em recursos cognitivos à disposição das práticas científicas. Um exemplo promissor pode ser encontrado nos trabalhos de Piccinini (2021; 2022). Sua abordagem acomoda todas as características associadas à cognição 4E. Não obstante, em linha similar à adotada por Clark (1997), ele defende a tese de que o caráter situado da cognição não acarreta incompatibilidade com a existência de mecanismos computacionais ou representacionais. Processos computacionais, inclusive os que exploram conteúdo representacional, são eles mesmos situados.

Dizer de um estado representacional que ele é situado significa dizer, entre outras coisas, que a determinação do seu conteúdo não implica a necessidade de representações adicionais. Em abordagens clássicas, representar a posição de um copo sobre a mesa implicava representar a mesa, o copo, quaisquer demais objetos presentes e, não bastasse, as devidas relações entre eles. Caso contrário, não seria possível atribuir ao estado representacional o papel de explicar como o organismo é capaz de antecipar possíveis desdobramentos (“se o copo for tocado de modo indevido, ele pode cair”). Em contraste, o conteúdo de representações situadas pode ser tal que leve em conta aspectos do corpo e do ambiente, resultando em conteúdos como “o copo pode ser alcançado com um leve girar da mão direita no sentido horário”.

Note-se que tal conteúdo só faz sentido para um dado organismo, com um dado corpo num dado estado e no interior de uma situação particular. Assim, o papel explicativo que esse conteúdo cumprirá na explicação de uma determinada capacidade é determinado de um modo situado. Ao contrário do que alguns sugerem, portanto, a ideia de representações situadas não mitiga o papel explicativo que representações agregam. Elas dizem respeito tão somente ao modo como representações são ou podem ser exploradas no interior de um dado sistema.

Na mesma linha, os mecanismos computacionais que encontramos no sistema neural estão longe de ser os mecanismos passivos que apenas reagem a *inputs* do ambiente a partir de seus estados internos. Pelo contrário, esses mecanismos se engajam em processos de aprendizagem ativa, acumulando informações de múltiplas fontes, tanto internas ao aparato neural (frequência e tempo de disparos neuronais, por exemplo) quanto externas (sinais visuais, auditivos, olfativos, etc.) Esse tipo de processo não apenas é compatível, mas dependente de o aparato cognitivo ser incorporado e integrado ao ambiente. Isso se dá não no sentido clássico segundo o qual corpo e ambiente são fontes de *inputs* e espaços para despejo de *outputs*, mas no sentido de que sistemas cognitivos exploram acoplamentos contínuos em tempo real.

Essa integração já acomoda uma dimensão de caráter enativista, pois ela permite entender como um organismo pode compreender o ambiente ao seu redor a partir do modo como o explora. Processos de aprendizagem ativa influenciam o corpo e ambiente ao mesmo tempo em que ambiente e corpo afetam o modo como os processos cognitivos se desenrolam. Por fim, processos de aprendizagem também podem envolver elementos afetivos, na medida em que o organismo dispõe de objetivos ou expectativas, certos *inputs* e/ou acoplamentos com o ambiente irão importar mais do que outros.

É importante salientar, contudo, que Piccinini (2021; 2022) não adota a teoria do conteúdo representacional de Cummins (1996). Ele endossa o uso de representações estruturais ancoradas no isomorfismo, e por isso sua abordagem também evita o problema duro do conteúdo. Contudo, Piccinini (2021; 2022) adota uma variante da teleosemântica que confunde fixação de alvos e determinação de conteúdo representacional.

Com efeito, seu tratamento do conteúdo representacional sofre dos problemas tratados nas seções anteriores. Em particular, o tratamento que Piccinini (2021; 2022) dá ao erro representacional ancora o que deveria ser um erro semântico em elementos não semânticos, tais como condições ambientais aquém das ideais ou mau funcionamento dos mecanismos envolvidos. Como vimos, isso torna sua abordagem mais vulnerável a argumentos pela trivialização do papel explicativo das representações. Felizmente, é possível ler o trabalho de Piccinini (2021; 2022) como tão somente uma descrição das estratégias que organismos utilizam para fixar alvos. Como a fixação de alvos é independente da determinação de conteúdo representacional, isso permite rejeitar a teoria do conteúdo representacional que Piccinini (2021; 2022) endossa em virtude da teoria de Cummins (1996) sem abrir mão daquilo que o projeto de Piccinini tem de melhor. O resultado desse amálgama é um bom exemplo do tipo de projeto que consideramos promissor: o que almeja acomodar as virtudes da cognição 4E e do cognitivismo sem carregar consigo seus vícios.

Evidentemente, é possível que essa compatibilidade repouse sobre diferentes concepções de cognição situada. Ser incorporado, integrado ou enativo pode significar coisas distintas. Mas ainda que existam, as consequências dessas divergências podem ser insuficientes para justificar uma clivagem. A distinção que Cummins (1996) faz entre alvo e conteúdo pode ser (novamente) útil para tornar esse ponto mais saliente: parte considerável do contraste entre representacionalistas e não representacionalistas sequer alcança a dimensão semântica.

Nesse sentido, o contraste é melhor compreendido como enfatizando diferentes modos pelos quais organismos são capazes de fixar em alvos, isto é., de exibirem alguma forma de intencionalidade. Se representações carregassem consigo o compromisso com abordagens não situadas da cognição, então sem dúvida o modo como sistemas representacionais conseguem fixar alvos (seja para representá-los, seja para explorá-los por meio de recursos não representacionais), seria incompatível. Essa restrição não se dá, contudo, sobre uma teoria da determinação do conteúdo representacional. Ela se dá sobre teses que associam a presença de representações ao tipo de mecanismo que as explora, uma confusão que Cummins (1996) nos ajuda a enxergar. Uma vez que sabemos a diferença entre o que é representar e o que é explorar, desdobra-se um cenário em que a revolução antirrepresentacionalista não é tão revolucionária assim, pois não implica (ainda que deseje) um completo abandono de mecanismos representacionais, mas tão somente de sua versão não situada. O poder explicativo que as *frameworks* tradicionais encontraram no conceito de representação não precisa ser descartado.

Conclusão

Neste artigo procuramos mostrar que há (pelo menos) uma noção de representação que não está atrelada ao “cognitivismo de velha guarda” e não implica o abandono de uma concepção da mente como *corporific ada, situada e enativa*. Procuramos também esclarecer que tipo de coisa pode ser uma representação e que papéis teóricos ela visa desempenhar, desfazendo os principais mal-entendidos associados a esta noção na filosofia da mente contemporânea. Finalmente, procuramos mostrar em que sentido a introdução de representações pode nos comprar uma dimensão explicativa extra que não está disponível fora deste vocabulário teórico. Se esta dimensão será de fato útil, é algo que deverá ser estabelecido empiricamente através de uma análise caso a caso, não podendo ser descartada *a priori* apenas em virtude das inclinações teóricas do pesquisador.

Nosso esforço nessa direção, mais uma vez, não é motivado por uma reabilitação do representacionalismo ou uma crítica ao enativismo. Nosso objetivo é apenas aliviar algumas preocupações dos enativistas acerca dos comprometimentos ontológicos e metodológicos associados ao vocabulário representacional, para que possam enfim reconhecer como possíveis aliados aqueles teóricos que também buscam investigar o caráter corporificado e situado da mente, ainda que o façam através de (uma certa noção de) representações mentais. Caso esse objetivo seja ao menos parcialmente realizado, poderemos ter a nosso dispor mais uma ferramenta teórica compartilhada, dissociada de uma vez por todas de suas origens computacionalistas e representacionalistas.

Referências

ANDERSEN, B. P.; MILLER, M.; VERVAEKE, J. Predictive processing and relevance realization: exploring convergent solutions to the frame problem. **Phenomenology and the Cognitive Sciences**, ago. 2022.

CAO, R. New Labels for Old Ideas: Predictive Processing and the Interpretation of Neural Signals. **Review of Philosophy and Psychology**, v. 11, n. 3, p. 517–46, ago. 2020.

CARVALHO, E. M. De; ROLLA, G. An Enactive-Ecological Approach to Information and Uncertainty. **Frontiers in Psychology**, v. 11, abr. 2020.

CLARK, A. **Being There: Putting Brain, Body, and World Together Again**. Cambridge: MIT Press, 1997.

CLARK, A. **Surfing Uncertainty**. Oxford: Oxford University Press, 2016.

CRYSTAL, J. D. Remembering the past and planning for the future in rats. **Behavioural Processes**, v. 93, p. 39–49, fev. 2013.

CUMMINS, R. **Representations, targets and attitudes**. MIT Press, 1996.

CUMMINS, R. Reply to Millikan. **Philosophy and Phenomenological Research**, v. 60, n. 1, p. 113, jan. 2000.

CUMMINS, R. et al. Representation and unexploited content. In: **The World in the Head**. Oxford, 2010. p. 120–33.

CUMMINS, R. How does it work versus “what are the laws?”: Two conceptions of psychological explanation. In: **The world in the head**. Oxford, 2010. p. 283–310.

DRETSKE, F. **Knowledge and the flow of information**. Cambridge: MIT Press, 1981.

DRETSKE, F. I. Misrepresentation. In: BOGDAN, Radu (Ed.). **Belief: Form, Content, and Function**. Oxford University Press, 1986. p. 17–36.

EGAN, F. A Deflationary Account of Mental Representation. In: SMORTCHKOVA, Joulia; DOLEGA, Krzysztof; SCHLICHT, Tobias (Eds.). **Mental Representations**. New York, USA: Oxford University Press, 2020.

FACCHIN, M. Structural representations do not meet the job description challenge. **Synthese**, v. 199, n. 3-4, p. 5479–5508, jan. 2021.

FODOR, J. A. **The Language of Thought**. Harvard University Press, 1980.

HAUGELAND, J. Truth and rule-following. In: **Having thought**. Cambridge: Harvard University Press, 1998. p. 305–61.

HEYES, C. **Cognitive Gadgets**. Harvard University Press, 2018.

HINTON, G. E. Learning multiple layers of representation. **Trends in Cognitive Sciences**, v. 11, n. 10, p. 428–34, out. 2007.

HURLEY, S. Perception and action: alternative views. **Synthese**, v. 129, n. 1, p. 3–40, 2001.

HUTTO, D.; MYIN, E. **Radicalizing enactivism: basic minds without content**. Cambridge, Mass: MIT Press, 2013.

MILLIKAN, R. G. **Language, Thought, and Other Biological Categories: New Foundations for Realism**. The MIT Press, 1987.

NEWEN, A.; VOSGERAU, G. Situated mental representations: why we need mental representations and how we should understand them. In: SMORTCHKOVA, Joulia; DOLEGA, Krzysztof; SCHLICHT, Tobias (Eds.). **What are mental representations?** Oxford University Press, 2020. p. 178–212.

NOË, A. **Action in Perception**. Cambridge: MIT Press, 2004.

PANOZ-BROWN, D. et al. Rats Remember Items in Context Using Episodic Memory. **Current Biology**, v. 26, n. 20, p. 2821–826, out. 2016.

PERLMAN, M. Pagan teleology: adaptational role and the philosophy of mind. In: ARIEW, André; CUMMINS, Robert; PERLMAN, Mark (Eds.). **Functions: new essays in the philosophy of psychology and biology**. Oxford University Press, 2002. p. 263–90.

PIANTADOSI, S. T.; JACOBS, R. A. Four Problems Solved by the Probabilistic Language

of Thought. **Current Directions in Psychological Science**, v. 25, n. 1, p. 54–9, fev. 2016.

PICCININI, G. **Neurocognitive Mechanisms: Explaining Biological Cognition**. Oxford University Press, 2021.

PICCININI, G. Situated Neural Representations: Solving the Problems of Content. **Frontiers in Neurobotics**, v. 16, abr. 2022.

RAMSEY, W. **Representation reconsidered**. Cambridge: Cambridge University Press, 2007.

ROLLA, G. **A mente enativa**. Porto Alegre: Editora Fi, 2021.

ROLLA, G. Por que não somos só o nosso cérebro: em defesa do enativismo. **TRANS/FORM/AÇÃO: Revista de Filosofia**, n. 46, p. 207–36, 2023.

SHEA, N. **Representation in Cognitive Science**. Oxford University Press, 2018.

STICH, S. **From folk-psychology to cognitive science: the case against belief**. MIT Press, 1983.

SWOYER, C. Structural representation and surrogate reasoning. **Synthese**, v. 87, n. 3, p. 449–508, jun. 1991.

VARELA, F. J.; ROSCH, E.; THOMPSON, E. T. **The Embodied Mind: cognitive science and human experience**. MIT Press, 1991.



