



Camila De Paoli Leporace

MACHINE LEARNING E A APRENDIZAGEM HUMANA
Uma Análise a Partir do Enativismo

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Educação do Departamento de Educação do Centro de Teologia e Ciências Humanas da PUC-Rio.

Orientador: Prof. Ralph Ings Bannell

Coorientadora: Prof. Maria Teresa
Ribeiro Pessoa

Rio de Janeiro
Janeiro de 2023



Camila de Paoli Leporace

MACHINE LEARNING E A APRENDIZAGEM HUMANA
Uma Análise a Partir do Enativismo

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Educação do Departamento de Educação do Centro de Teologia e Ciências Humanas da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Ralph Ings Bannell

Orientador

Departamento de Educação - PUC-Rio

Prof^a. Maria Teresa Ribeiro Pessoa

Coorientadora

Universidade de Coimbra-PT

Prof^a. Maria Piedade Simões Santana Pessoa Vaz Rebelo

Universidade de Coimbra-PT

Prof. Giovanni Rolla

UFBA

Prof. Edgar de Brito Lyra Neto

Departamento de Filosofia - PUC-Rio

Prof. Thiago Leite Cabrera Pereira Rosa

Departamento de Educação - PUC-Rio

Rio de Janeiro, 6 de janeiro de 2023.

Todos os direitos reservados. É proibida a reprodução total ou parcial sem autorização da Universidade, do Autor e dos Orientadores.

Camila De Paoli Leporace

Mestre em Educação pela Pontifícia Universidade Católica do Rio de Janeiro/PUC-Rio (2019), graduou-se em Jornalismo pela mesma instituição em 2005. Tem experiência em comunicação para variadas mídias, educação a distância e tecnologias educacionais digitais. É integrante do GEPFE – Grupo de Estudos e Pesquisa em Filosofia e Educação, coordenado pelo professor Ralph Ings Bannell.

Ficha Catalográfica

Leporace, Camila De Paoli

Machine learning e a aprendizagem humana: uma análise a partir do enativismo / Camila De Paoli Leporace; orientador: Ralph Ings Bannell; coorientadora: Maria Teresa Ribeiro Pessoa – 2023.

217 f.; 30 cm

Tese (doutorado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Educação, 2023.
Inclui bibliografia

1. Educação – Teses. 2. Machine learning. 3. Aprendizagem. 4. Enativismo. 5. Cognição. 6. Autonomia. I. Bannell, Ralph Ings. II. Pessoa, Maria Teresa Ribeiro. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Educação. IV. Título.

CDD: 370

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O período em Portugal aconteceu no âmbito do projeto CAPES PrInt.

Agradeço à PUC-Rio e à Universidade de Coimbra pelo apoio que possibilitou a realização do doutorado em regime de cotutela.

Aos professores da PUC-Rio e da Universidade de Coimbra, pelo conhecimento compartilhado e pelo suporte durante todo o percurso.

Ao meu orientador na PUC-Rio desde o mestrado, professor Ralph, pela parceria e pelas trocas constantes, por compartilhar seu conhecimento comigo e me incentivar a ir mais longe.

À minha orientadora na Universidade de Coimbra, professora Maria Teresa, pelo acolhimento, incentivo e apoio.

Ao professor Eduardo por me apresentar a Universidade de Coimbra e pelas parcerias desenvolvidas.

Aos professores que fizeram parte das minhas bancas – ou júris – da primeira qualificação à defesa da tese.

Aos pareceristas anônimos das publicações científicas que avaliaram trabalhos meus e assim também me ajudaram a aprimorá-los.

A aqueles que me fizeram perguntas instigantes e trocaram tanto comigo, quando ministrei aulas e apresentações.

Aos colegas da PUC-Rio e da Universidade de Coimbra.

Na PUC, agradeço especialmente aos parceiros do GEPFE – Grupo de Pesquisa em Filosofia e Educação da PUC-Rio, liderado pelo prof. Ralph Bannell.

Aos amigos que se fizeram presentes e foram acolhedores sempre, especialmente Luís, Adriana, Lorena, Bruna, Isa, Mauricio, Fernanda, Aline, Maria Carolina, Vitoria e Margareth.

Aos amigos que Portugal me deu, especialmente Adriana, Juju & Lali, Yasmin – minha “irmã” mais nova, Fabi, Nay, Ana, Neli, Alex, Pedro, Pedro & Mel e Raoni.

Ao Seixas, pela amizade, pelo apoio e o incentivo.

À minha família, especialmente aos meus pais, Sandra e Zeca, e aos meus padrinhos, Mary e Arnaldo, terei que agradecer infinitamente – e ainda assim não terei dado conta.

As Merleau, ao Yoda, ao Chicote e à Pretinha pelas trocas de energia essenciais e por continuarem me lembrando que os animais são essenciais em nosso cotidiano.

Resumo

Leporace, Camila De Paoli; Bannell, Ralph Ings.; Pessoa, Maria Teresa (Orientador e Coorientadora) — “**Machine learning e a aprendizagem humana**: uma análise a partir do enativismo”. Rio de Janeiro, 2023. 217p. Tese de Doutorado – Departamento de Educação, Pontifícia Universidade Católica do Rio de Janeiro, e Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Portugal.

Situada no campo da filosofia da educação, a tese dialoga também com o campo das tecnologias educacionais. O trabalho busca uma compreensão filosófica dos impactos da aprendizagem de máquina ou machine learning na educação. Para isso, dedica-se aos pressupostos subjacentes à aprendizagem de máquina em articulação com os pressupostos subjacentes à concepção de aprendizagem humana que descende do enativismo. Defende-se que a chegada da aprendizagem de máquina na educação encontra um campo em que ainda predomina o paradigma cognitivista, o qual é bastante profícuo para que germinem as tecnologias baseadas em dados e redes neurais. Avança-se para demonstrar que esse paradigma, no entanto, vem sendo desafiado por outras abordagens de pesquisa que se dedicam à mente humana, dentre as quais se destaca o enativismo. São explicitadas as bases teóricas fundamentais do enativismo, e como elas se desdobram em pressupostos para uma aprendizagem humana que é corporificada e essencialmente orientada ao acoplamento do ser com o mundo e com os outros agentes. É dedicada atenção especial aos impactos da aprendizagem de máquina na autonomia do cognoscente, a qual, sob a perspectiva do enativismo, somente pode existir e se manter nas trocas com o meio e com aqueles que habitam e formam esse ambiente. Demonstra-se que, para que as tecnologias algorítmicas sejam adequadas a uma concepção de cognição e de aprendizagem enativista, é preciso buscar um caminho de valorização ainda maior do corpo na aprendizagem, bem como da intersubjetividade, uma vez que as relações entre os agentes cognitivos não são concebidas como articulações opcionais, mas como um elemento que está no cerne da atividade cognitiva humana e do qual essa atividade emerge.

Palavras-Chave

Machine learning; aprendizagem; enativismo; cognição; autonomia

Abstract

Leporace, Camila De Paoli; Bannell, Ralph Ings.; Pessoa, Maria Teresa (Advisor and Co-advisor) — “**Machine learning and human learning: an enactivist analysis**”. Rio de Janeiro, 2023. 217p. Tese de Doutorado – Departamento de Educação, Pontifícia Universidade Católica do Rio de Janeiro, e Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Portugal.

This work is situated in the field of philosophy of education, and also relates to the field of educational technologies. The thesis seeks a philosophical understanding of the impacts of machine learning in education. To do so, it addresses the assumptions underlying machine learning in conjunction with the premises underlying the conception of human learning that derive from enactivism. It is argued that the arrival of machine learning in education found fertile ground in which the cognitivist paradigm still prevails, a situation that is rather fruitful for technologies based on data and neural networks to thrive. The thesis demonstrates that this paradigm, however, has been challenged by other research approaches that are dedicated to the human mind, among which enactivism is emphasized. The fundamental theoretical underpinnings of autopoietic enactivism are explained, as well as how they unfold in assumptions for a notion of human learning that is embodied and essentially oriented towards the coupling of the human organism with the world and with other agents. Particular attention is drawn to the impacts of machine learning on the autonomy of the cognizer, which, from the perspective of enactivism, can only exist and be maintained in exchanges with the environment and with those who inhabit and shape this environment. It is shown that for algorithmic technologies to be suited to an enactivist conception of cognition and learning a greater appreciation of the body in learning is necessary, as well as intersubjectivity, since the connections between cognitive agents are not conceived as optional articulations, but as an element that is at the core of human cognitive activity and from which this activity emerges.

Keywords

Machine learning; learning; enactivism; cognition; autonomy

Sumário

1 Introdução.....	15
2 Máquinas para aprender, máquinas para ensinar: uma perspectiva da aprendizagem automatizada no contexto da educação.....	29
2.1 Breve definição de machine learning	30
2.1.1 Big Data	34
2.2 Inteligência artificial na educação	37
2.3 Programando o ensino.....	41
2.3.1 Ensino adaptativo: primórdios.....	44
2.3.2 Derrocada das (primeiras) máquinas de ensinar	49
2.4 Máquinas de ensinar ontem e hoje.....	53
2.5 Estendendo a discussão	60
2.6 Conclusão do capítulo	63
3 Tecnologias “cerebrais”	65
3.1 Tecnologias do cérebro e para o cérebro.....	66
3.2 A eclosão do cognitivismo e da inteligência artificial	75
3.3 Problemas do paradigma cognitivista evidenciados pela IA.....	79
3.3.1 O que os computadores <i>não</i> fazem diz muito sobre os humanos.....	81
3.4 Estendendo a discussão: a vida que não cabe em um tutorial	85
3.4.1 O contexto importa: o problema da relevância	86
3.5 Conclusão do capítulo	93
4 A cognição enativa: a mente e(é) a vida	97
4.1 (N)o caminho do imprevisível: percepção e(é) ação.....	101
4.2 Sobre a vertente do enativismo aqui adotada.....	103
4.3 A identidade do ser cognoscente: onde começa a cognição?	106
4.4 Autonomia.....	111
4.5 Emergência.....	117
4.6 <i>Sense-making</i>	118
4.6.1 Uma inversão relevante na compreensão tradicional da cognição.....	123
4.7 <i>Linguistic bodies</i> : a corporificação e a linguagem no enativismo.....	126
4.8 Conclusão do capítulo	133
5 Enativismo, autonomia e machine learning.....	135
5.1 O conceito de <i>participatory sense-making</i>	140

5.2 <i>Participatory sense-making e machine learning</i>	149
5.3 Ao lidar com sistemas de machine learning, seguimos sendo <i>sense-makers</i> autônomos?	152
5.4 Estendendo a discussão: modelos matemáticos, fragmentos do mundo	162
5.5 Conclusão do capítulo	165
6 Abrindo a caixa preta	168
6.1 Síntese de uma concepção enativista para a aprendizagem	171
6.2 Da centralidade do corpo na aprendizagem.....	173
6.2.1 Corpo, transformação e percepção da aprendizagem	180
6.3 Da interconexão entre mentes na busca de alternativas ao problema da autonomia.....	185
6.3.1 Em busca de equilíbrio	189
6.4 Conclusão do capítulo	195
7 Considerações finais	199
8 Referências Bibliográficas	204
Apêndice	213

Lista de figuras

Figura 1 - Pilares da pesquisa	22
Figura 2 – Continuidade vida-mente	125

The future of artificial intelligence (AI) in education — as with any aspect of the future — is uncertain, unpredictable, and essentially unknowable.

(Selwyn, 2022)

Prefácio

Certa vez, em uma aula que ministrei, uma aluna perguntou algo como: “Mas por que, se a ideia é reduzirmos o nosso esforço cognitivo para fazer as coisas (ela se referia a um conceito presente na tese da mente estendida), às vezes escolhemos os caminhos mais longos para ir de um lugar a outro, por exemplo?”. Dialogando com ela, respondi algo como: “O ser humano não preza apenas pelo que é objetivo, mas pela sua experiência e pela criação de ambientes e hábitos que lhes sejam favoráveis. Talvez variar o caminho seja uma estratégia que sabemos que poderá nos ajudar a ser mais criativos ou aliviar o estresse, por exemplo, se for um caminho bonito, se for interessante, se nos mostrar algo mais...”. Tudo o que fazemos para descobrir o caminho *certo* pode ser posto de lado momentaneamente se decidirmos pegar uma rota só pela aventura, por ter uma vista diferente ou porque podemos ouvir passarinhos ou encontrar um amigo no caminho. Senão isso, por que viajarmos? Por que descobriremos novos espaços, sabores, lugares e pessoas? Buscar isso não é inútil e não nos torna menos humanos, muito pelo contrário! É justamente esse tipo de decisão que os algoritmos não podem tomar por nós: aquelas que fogem à simples resolução de tarefas. Seres humanos não apenas resolvem problemas e cumprem metas. Não existimos apenas para isso. Resgatando o que Maturana e Varela colocaram e que continua sendo imperial ao enativismo, *o caminho se faz ao caminhar*. Faz parte do sentido de *ser* humano não ter tudo já traçado de antemão: descobre-se o que fazer, e como fazer, na medida em que se movimenta, arrisca, experimenta.

Este trabalho é mais sobre a aprendizagem humana do que sobre a aprendizagem de máquina, mesmo que se debruce sobre uma compreensão filosófica do sentido da aprendizagem de máquina na educação. Ao buscar as premissas da aprendizagem de máquina, e analisá-las sob a perspectiva do enativismo, fica em evidência a aprendizagem humana e aquilo que pode emergir quando se associa a aprendizagem de máquina à aprendizagem humana. O trabalho pode apresentar, em alguns momentos, um tom talvez negativo com relação às tecnologias digitais baseadas em algoritmos, especificamente aquelas fundamentadas na chamada aprendizagem de máquina ou machine learning. Trata-se, afinal, de um trabalho analítico, que busca analisar os possíveis impactos de tais tecnologias, partindo da perspectiva ampla da relação humana com esses recursos.

Mas não seria justamente esse o papel dos pesquisadores do campo? Isto é, não seríamos nós responsáveis por levantar questões complexas, indicar direções de análise que outros campos possivelmente não estão observando, aprofundar determinadas temáticas buscando os vários lados que elas fazem emergir?

Ao conduzir essas análises, meu intuito aqui não é o de me posicionar contra o uso de tecnologias educacionais baseadas em inteligência artificial. Muito pelo contrário, aliás. Sempre trabalhei com elas, as considero parte do jogo, não apenas da educação, mas daquele que permeia a nossa sociedade. Mas foi do meu próprio trabalho com tecnologias dessa natureza que surgiram diversas preocupações a respeito de como estamos evoluindo, uma vez imersos numa imensa *algoritmosfera*. Somos nós que alimentamos esse ambiente com nossos dados, e eles nos alimentam com mais informações, e assim sucessivamente. A minha preocupação, bem como a minha crítica, volta-se muito mais para esse complexo mundo algorítmico, que nos desafia com tantas questões que vão se abrindo e que são ainda novas para nós, do que propriamente com tecnologias usadas de maneira pontual, com finalidades específicas, para ensinar determinados conteúdos ou cumprir com determinadas tarefas.

Uma questão é que as plataformas adaptativas de uso escolar, por exemplo, podem por vezes ser consideradas bem-sucedidas quando aplicadas para ensinar um determinado tópico, uma certa disciplina; porém, a ideia das plataformas de aprendizagem acabou se disseminando por todo o universo web, com outras repercussões. Estamos agora ligados a plataformas baseadas em machine learning quando fazemos os mais diversos tipos de atividade na rede. Então, se podemos dizer que isso será cada vez mais comum e que de fato podemos aprender continuamente sobre o mundo enquanto usamos a web, precisamos também olhar para a relação que estabelecemos com esses recursos e para a nossa capacidade de utilizá-los da melhor maneira para obter conhecimento.

Toda a análise crítica desta pesquisa parte de um pressuposto: tecnologias digitais, sejam elas baseadas em inteligência artificial ou não, são complementares à cognição humana. Na minha pesquisa de mestrado, que finalizei em 2018, eu concluí, a partir da tese da mente estendida desenvolvida por Andy Clark, que o nosso foco como sociedade digital deve ser na complementação e não na substituição. Principalmente quando se trata de educação. Tendo esse ponto de partida, vale dizer o que esta pesquisa *não* é: eu *não* desenvolvi uma pesquisa de

doutorado para procurar entender se podemos ou devemos ser substituídos por máquinas. Desenvolvi-a para aprofundar a nossa relação com essas tecnologias, descobrir o que significa dizer que sistemas "aprendem" (sobre nós) e o que efetivamente podemos aprender com eles, ou sobre eles, em nosso dia a dia de uso constante da www, que se torna cada vez mais algorítmica. Principalmente, a pesquisa direciona-se para uma concepção de aprendizagem que vem surgindo no horizonte e que nos desafia acerca de como podemos pensar o nosso equilíbrio com as tecnologias digitais. Equilíbrio, aliás, é uma palavra interessante: precisamos das tecnologias e elas, de nós e de nossos dados; então, se isto é possível, como é que podemos obter um resultado interessante desse *acoplamento* com tais recursos?

É buscando analisar o que acontece quando lidamos com dispositivos baseados em algoritmos que podemos aprender a operar melhor com eles, mantendo positivas as trocas que temos com esses sistemas. Espero contribuir nesse sentido, trazendo um enfoque a partir de uma perspectiva ainda pouco explorada na educação, e assim incluindo ingredientes mais novos neste já tão desafiador movimento. Também pretendo contribuir, de maneira especial, com os professores e educadores que assumem diversos papéis e que se sentem desafiados com este mundo algorítmico em que vivemos. Existem elementos que são tipicamente humanos, e – até o momento pelo menos – são unicamente humanos. Eles são bastante explorados na análise que esta tese desenvolve. Espero que o trabalho some a outros no sentido de contribuir para trazer o enativismo para o campo da educação como uma perspectiva para pensar a aprendizagem, em específico a presença da aprendizagem de máquina na aprendizagem humana.

1

Introdução

Situada no campo da filosofia da educação, esta tese propõe uma contraposição entre a concepção de aprendizagem presente nos sistemas de aprendizagem de máquina, ou machine learning – vertente da ciência da computação em destaque no campo da inteligência artificial hoje – e o conceito de aprendizagem que pode ser apreendido a partir da abordagem enativista para a cognição. Parto da hipótese de que a “aprendizagem”, no contexto de sistemas baseados em previsões e estatísticas (os quais vêm sendo trazidos para o campo da educação para participar do cotidiano de aulas e avaliações de alunos e professores e já estão sendo usados largamente em sistemas online com os quais lidamos cotidianamente), desconsidera aspectos fundamentais das faculdades cognitivas humanas, uma vez que se adote uma perspectiva enativista. A tese é desenvolvida de modo a investigar essa hipótese. No decorrer da investigação, foram constatadas razões para considerá-la verdadeira.

A abordagem enativista apresenta uma concepção da mente que, essencialmente, diverge da concepção cognitivista – surgida junto ao nascimento da ciência cognitiva enquanto área de pesquisa na década de 1950 – e mesmo da abordagem conexionista, que pode ser considerada uma variação do cognitivismo

tradicional ou clássico (ver, por exemplo, Clark, 2014). Em busca de uma perspectiva alternativa para o estudo da cognição, é adotada, nesta tese, a concepção de enativismo tal como desenvolvida principalmente por Di Paolo e colegas, especialmente Di Paolo, Buhrmann e Barandiaran (2014); Di Paolo, Cuffari e De Jaegher (2018); De Jaegher e Di Paolo (2007, 2008) e Colombetti (2014). A contribuição do filósofo Hubert Dreyfus (1975, 1992, 2009 e outras) é fundamental como inspiração e horizonte para este trabalho. Os aportes de Thompson (2007),

Maturana e Varela (1980, 2019) Johnson (2007, 2017, 2018) e outros são considerados de maneira a complementar as proposições dos autores supracitados.

Antes de começar esta pesquisa, uma pergunta vinha-me à mente: por que os sistemas de machine learning são chamados assim? Pode-se dizer que máquinas realmente *aprendem*? Aprendem como seres humanos? Para procurar possíveis respostas a essa questão, é preciso compreender a concepção de aprendizagem de que se está falando. De antemão, pode-se dizer que aprendizagem, em machine

learning, é algo bastante restrito, pois se refere a estatística – isto é, previsões feitas a partir de dados. Por mais que as máquinas possam avançar, uma vez que a leitura e a combinação de dados geram novos dados para serem usados nas previsões, elas operam em sistemas fechados, pré-formatados, que seguem determinados objetivos traçados. Isso faz com que sejam “cegas” para o mundo – num sentido que demonstrarei no decorrer deste trabalho. Sistemas de machine learning acabam sendo “caixas pretas” em, pelo menos, dois sentidos: primeiro, porque geram imensas dúvidas a respeito de como funcionam, deixando por vezes confusas as pessoas que usam os sistemas e que são afetadas por eles – não raro, pessoas que exercem diversas funções, entre elas professores, se questionam se suas profissões poderão ser substituídas por sistemas artificiais num futuro próximo; segundo, porque são como caixas herméticas inconscientes, incorpóreas, incapazes de ter qualquer experiência do mundo, de fato. De certo modo, a mente humana também é uma caixa preta, ainda. Não no sentido como se compreendia à época em que o behaviorismo era a teoria mais forte na pesquisa, mas no sentido de que ainda há muita dúvida sobre como ela funciona – e até sobre o que a mente é e como é constituída.

Seria possível *aprender* sem perceber ou sem experimentar o mundo?

A educação já sabe que não, não é possível. No século XX, os psicólogos Vygostky, Luria e Leontiev questionaram o modo como a psicologia propunha o desenvolvimento das capacidades cognitivas. Uma das contribuições mais importantes desses investigadores foi a noção de que o desenvolvimento cognitivo é, ao mesmo tempo, um processo interpessoal e intrapessoal, isto é, a mente não se limita a funções internas e individuais. Vygotsky, em especial, “não concordava que o cérebro seria o *locus* isolado do conhecimento” e em 1932 considerou as emoções como funções psicológicas superiores (Bannell et al., 2017, p. 59).

Teorias construtivistas sociais há anos demonstram a importância das trocas entre o indivíduo e o meio, valorizando as interações culturais e sociais e assim buscando alternativas ao dualismo cartesiano (mente/cérebro). Como indica Pessoa (2002),

O construtivismo é considerado como uma teoria da aprendizagem e como uma teoria sobre conhecimento (Fosnot, 1999a, 9), ou como uma forma de pensar sobre os mesmos (MacKinnon & Scarff-Seatter, 1997, 51), com raízes, como referimos, no domínio da psicologia e da filosofia (Dalgarno, 1996; Doolittle & Camp, 1999) e tem vindo a desenvolver-se como paradigma alternativo ao objectivismo rejeitando,

assim, o seu princípio fundamental, isto é, de que o conhecimento poderá ser representado de forma objectiva e correcta para além do sujeito que conhece.

O construtivismo tem como princípio fundamental, que vem sendo partilhado pelos diversos construtivismos, o primado absoluto do sujeito na construção do conhecimento na medida do valor que lhe atribui. Como realça Le Moigne, o conhecimento implica um sujeito *connaissant* e não tem valor ou significado para além dele (1999, 67).

Esta construção do conhecimento assenta, então, na experiência do próprio sujeito da realidade. Assim, os significados e as perspectivas acerca dos acontecimentos ou fenómenos são construídos pelos sujeitos com base nas suas experiências da realidade. (Pessoa, 2002, p. 134).

Apesar de a educação levar em conta, há bastante tempo, o protagonismo do sujeito e os impactos sociais e culturais, o destaque reservado ao corpo e à afetividade na aprendizagem poderia ser ampliado radicalmente com a perspectiva enativista. A ideia dessa abordagem, como neste trabalho se pretende mostrar, é colocar tais elementos no centro, de modo que a partir deles emergja a mente humana, com os seus processos cognitivos. Para isso, são indicadas pelos enativistas novas maneiras, mais abrangentes, de se conceber o corpo e a afetividade. A educação precisa olhar para essas propostas contemporâneas, que enraízam a mente no corpo, não para descartar o trabalho já desenvolvido por outros como Vygostky, Luria e Leontiev, mas justamente para dar continuidade ao trabalho de transpor para o campo da educação as pesquisas avançadas sobre a cognição humana. Um dos riscos de não fazer isso, e que aqui cabe destacar, é o de permitir que o campo da educação seja permeado por tecnologias digitais cérebro-centradas, como aqui defendo que são os sistemas algorítmicos, sem que esse movimento seja acompanhado de uma reflexão crítica que se traduza em práticas de aprendizagem para além do processamento cerebral.

Bannell et al. abordam a importância da interlocução entre a educação e as teorias cognitivas, especialmente no que diz respeito às tecnologias digitais no âmbito educacional:

Mesmo que tenha a aprendizagem como foco, a maioria das ações de inserção da tecnologia nas escolas desconsidera os estudos sobre cognição, que levam em conta as mudanças culturais produzidas pela onipresença das tecnologias digitais nas sociedades contemporâneas. De modo geral, a formação dos professores e os projetos e modelos educacionais associados ao uso de computadores, internet e dispositivos móveis na educação formal seguem ancorados em conceitos e pressupostos sobre cognição e aprendizagem formulados em contextos históricos muito distintos deste em que estamos vivendo e no qual estão sendo formadas as novas gerações. Para realizar mudanças significativas nas práticas educativas, de modo que a escola possa atender às demandas sociais do século XXI, precisamos

rever e atualizar conceitos e teorias sobre a cognição e sobre como os seres humanos aprendem. (Bannell et al., 2017, p. 57).

Este trabalho tem como objetivo contribuir neste sentido que os autores indicam acima. A minha proposta é colaborar para que a educação adquira meios para avaliar os impactos da vertente da inteligência artificial chamada de aprendizagem de máquina – que tem se traduzido num recurso tecnológico cada vez mais presente em nossas vidas – a partir da visão enativista para a cognição. Procuro fazer isso apresentando a concepção de aprendizagem que pode emergir do enativismo, de modo a contrastá-la com a concepção de aprendizagem presente em machine learning. Algumas questões impulsionadoras são as seguintes: qual o conceito de aprendizagem subjacente à aprendizagem de máquina? Quais os riscos de esse conceito ser confundido com a maneira como a aprendizagem humana acontece? Se a aprendizagem humana é diferente da aprendizagem de máquina, por que é diferente? Quais os elementos que marcam a aprendizagem humana? Essas são perguntas que passam pelos pressupostos, investigam conceitos e, ao fazer isso, podem indicar caminhos para elaborarmos qual pode ser o sentido da presença de tecnologias avançadas na aprendizagem.

Com a contribuição que procuro dar com os resultados desta pesquisa, desejo ir à base daquilo que se pode chamar de tecnologia cognitiva; trata-se de tecnologias que surgem a partir das capacidades cognitivas humanas, seja para substituí-las, seja para complementá-las ou aproveitá-las de alguma maneira; acredito que, ao investigar os elementos envolvidos na cognição, podemos compreender melhor o sentido dos sistemas aos quais se atribui a capacidade de aprender. Afinal, é preciso compreender o que a aprendizagem é ou pode ser, se desejamos integrar a esses processos sistemas capazes de contribuir com eles de alguma maneira (ou para percebermos que não precisamos deles para toda e qualquer situação e então decidirmos o que fazer e como).

É olhando para nossas capacidades cognitivas, para a emergência daquilo que entendemos por cognição, raciocínio, memória, inteligência, e ainda buscando interseções entre esses processos e as emoções, o corpo e as nossas relações com as outras pessoas que conseguimos compreender melhor o que sistemas artificiais poderiam fazer por nós, para nós e junto conosco. Não me dedico profundamente às questões que tangem à democracia e à ética que vêm emergindo com o machine learning e a ciência de dados, e que compõem um campo próprio de conhecimento,

apesar de reconhecer seu entrelaçamento com algumas das questões da tese e, claro, de notar que endereçá-las é de profunda importância. Mas acredito que, ainda que por outro enfoque, esta pesquisa poderá contribuir a seu modo com as análises sobre como poderemos encaminhar essas questões enquanto sociedade e enquanto educadores.

Contexto

Quando a ciência cognitiva começou a se desenhar como campo de pesquisa, nos anos de 1950, a abordagem predominante para a mente humana era a que se ajustava à metáfora computacional (da mente como um software rodando no cérebro-*hardware*), a qual culminou no programa de pesquisa que ficou conhecido como cognitivismo. A abordagem se fundamenta essencialmente no processamento de informações, mediado por símbolos ou representações mentais¹. Hoje, apesar de essa perspectiva ainda ser dotada de bastante destaque, aparecendo de forma extensiva como fundamento de pesquisas empíricas no campo da psicologia, linguística, neurociência e outros, ela já deixou de reinar absoluta; não apenas por conta do avanço do connexionismo – que, como já mencionado, poderia ser considerado uma variação do cognitivismo (Clark, 2014), mas pelo aporte de teorias subsequentes. Na educação, como mencionado, a predominância é das teorias construtivistas sociais.

A pesquisa em cognição tem sido permeada também por abordagens que focam, de diversas maneiras, no ambiente e na cultura. Elas vieram a ser conhecidas especialmente nos últimos 40 anos como abordagens *situadas* à mente humana, compondo um conjunto chamado de *4Es* – *Embedded*² (Situada); *Embodied* (Corporificada)³; *Extended* (Estendida) e *Enactive* (Enativa) (Newen, De Bruin e Gallagher, 2018). Ainda que apresentem diferenças conceituais sensíveis e variadas, essas vertentes fazem propostas que contam com um ponto comum: a ideia de que a cognição está para além do cérebro. Mais do que isso, trata-se de abordagens que deslocam o foco no pensamento abstrato para uma perspectiva que dá foco à ação corporificada. Elas promovem uma aproximação entre ação,

¹ Trabalhei estes tópicos, bem como um histórico da ciência cognitiva pelo viés da filosofia, em minha dissertação de mestrado (Leporace, 2019).

² Também referida constantemente como Embutida.

³ Também referida constantemente como Incorporada.

percepção e as emoções, assumindo o corpo enraizado no ambiente como locus fundamental de toda a atividade cognitiva. Nas citações a seguir, pesquisadores do enativismo indicam o caminho alternativo que vem sendo trilhado por essas investigações:

De forma constante, como o movimento das placas tectônicas, a ciência cognitiva tem vivido uma passagem do cognitivismo abstrato para abordagens mais dinâmicas, corporificadas, embutidas, estendidas e situadas (Di Paolo, Buhrmann & Barandiaran, 2017, p. 13)⁴.

Um olhar cuidadoso para aquilo que sabemos sobre nossos corpos, sua biologia, e a maneira como se organizam em poderes e sensibilidades, um olhar para a forma como experimentamos o mundo como seres situados em relações complexas com outros conta uma história bastante diferente (Di Paolo, Buhrmann & Barandiaran, 2017, p. 2)⁵.

A “história” que os pesquisadores enativistas mencionam na segunda citação destas acima envolve o corpo em suas variadas dimensões, a relação direta que esse corpo tem com o mundo (físico, cultural e social) e uma série de outros aspectos que daí decorrem. Para isso, trazem para o primeiro plano o corpo e sua atividade, a qual se desenrola em acoplamento com o mundo, de uma maneira sistêmica e em caráter não opcional; isto é, na sua presença no mundo é que esse corpo faz emergir toda a sua capacidade cognitiva, sendo, portanto, mandatório considerar o corpo como parte constitutiva da mente.

O *E* do enativismo tem se mostrado particularmente empenhado em trazer o corpo para o centro desse contexto, integrando também as emoções como aspectos constitutivos da cognição (é até mais preciso falar em afetividade do que em emoções como atributo essencial da cognição sob o ponto de vista enativista, como demonstrarei no decorrer do trabalho). Trata-se de um trabalho em curso e ainda com muito a desenvolver, como o trecho a seguir confirma:

O jogo virou completamente? Estamos agora vivendo a era de ouro da ciência cognitiva corporificada? A resposta é um sonoro ‘não’. (...) [M]uito do trabalho em

⁴ Todas as traduções são da autora. No original: *Steadily as the movement of tectonic plates, cognitive science has been shifting from abstract cognitivism to more dynamical, embodied, embedded, extended, and situated approaches.* (Di Paolo, Buhrmann & Barandiaran, 2017, p. 13)

⁵ No original: *A careful look at what we know about our bodies, about their biology, and the way they organize themselves into powers and sensitivities, a look at the way we experience the world as situated creatures in complex relations to other creatures tells a rather different story.* (Di Paolo, Buhrmann & Barandiaran, 2017, p. 2)

ciência cognitiva permanece dentro da hipótese da mente como uma máquina de processar informações (Di Paolo, Buhrmann & Barandiaran, 2017, p. 13)⁶.

No campo da ciência cognitiva, porém, ainda persistem abordagens representacionais e computacionais à mente humana, ou ao menos há ainda muito a ser desenvolvido na elaboração de propostas alternativas ou complementares a essas. Por outro lado, sobre a educação ainda paira uma sombra do sujeito moderno cartesiano, o sujeito desprendido (Bannell et. al, 2016; Bannell, 2019). Isto é, por mais que se admita que a mente depende do contexto social e histórico, o qual é importante para o sujeito; e que se admita que o corpo é essencial para a atividade cognitiva, não se pode desconsiderar a tendência ainda forte de limitar a mente ao cérebro. Essa tendência é forte especialmente quando se trata de tecnologias digitais. Há ainda um grande esforço de pesquisa voltado para o cérebro que tem como pressuposto ser possível, a partir do foco exclusivo em seu funcionamento, se chegar à compreensão total da atividade cognitiva (e, conseqüentemente, da aprendizagem). Ao encontro dessas pesquisas, tem se popularizado a aproximação entre a neurociência e a educação, por vezes referida como neuropedagogia ou neuroeducação.

Temos, então, um cenário em que convivem, de um lado, o crescimento da aprendizagem de máquina na inteligência artificial (IA) condizente com o enfoque na neurociência no campo da educação e, de outro, o incremento da pesquisa com foco no corpo e nas emoções, no cenário da ciência cognitiva. É importante analisar como nós, educadores, podemos nos posicionar nesse cenário.

Voltemos, neste momento, a nossa atenção para a IA. Já faz um tempo que o campo da IA tem colaborado, ainda que sem querer, para a constatação de que a mente humana não é apenas o cérebro. A educação, por sua vez, parece ter muito a ganhar a partir da observação daquilo que esse campo tem pesquisado. Afinal, os desafios da IA, especialmente na forma da aprendizagem de máquina/machine learning, envolvem questões intrínsecas, fundamentais, e que são semelhantes a algumas das preocupações centrais da educação. Como oportunamente colocaram os autores a seguir,

⁶ No original: *Have the tables turned completely? Are we now living in the golden age of embodied cognitive science? The answer must be a qualified 'no'. (...) [M]uch of the work in cognitive science remains well within the assumption of the mind as an information processing machine.* (Di Paolo, Buhrmann & Barandiaran, 2017, p. 13)

A inteligência artificial (IA) é um assunto que está intimamente conectado à educação. A IA se preocupa em fazer com que máquinas leiam, raciocinem, se expressem, façam generalizações e aprendam. No fim das contas, essas são as questões da educação”⁷. (Schank & Edelson, 1989, p. 3)

Considero esse um fato significativo para uma discussão teórica como a que aqui apresento. Afinal, a partir de tal constatação, pode-se analisar a aprendizagem de máquina ou machine learning não como algo estritamente técnico que, oriundo de outro campo (a ciência da computação, a programação), se deslocaria para o campo da educação para ser apropriado como uma tecnologia educacional, isto é, como um recurso a ser usado (em sala de aula ou fora dela) para ensinar ou aprender. Em vez disso, machine learning abre um campo de pesquisa que, desde a sua própria gênese, e em sua essência, compartilha com a educação questões que precisam ser investigadas. Trata-se de investigações de base para ambas as áreas.

Há alguns pilares essenciais à pesquisa que me propus a fazer: machine learning/IA e seus pressupostos e a aprendizagem humana e seus pressupostos, posicionados sobre um terceiro: a compreensão da cognição humana, sob as lentes do enativismo.

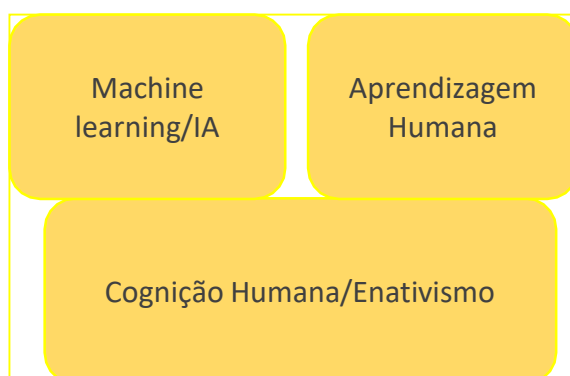


Figura 1 – Pilares da pesquisa (imagem da autora)

A aprendizagem de máquina ou machine learning⁸ traz em sua própria concepção uma determinada perspectiva da aprendizagem – e isso acontece em paralelo à disseminação de novas teses que investigam a mente humana, trazendo abordagens alternativas para compreendermos a cognição. A pesquisa sobre a

⁷ No original: *Artificial Intelligence (AI) is a subject that is intimately bound up with education. AI concerns itself with getting machines to read, to reason, to express themselves, to make generalizations, and to learn. These issues are the stuff of education after all* (Schank & Edelson, 1989, p. 3–4).

⁸ Para simplificar, por vezes refiro-me a machine learning como ML.

cognição pode abrir um caminho para a pesquisa sobre como a aprendizagem acontece; por outro lado, sem um diálogo entre essas áreas, algo se perde, uma vez que a compreensão dos processos cognitivos deveria informar a maneira como se entende a aprendizagem e, conseqüentemente, as estratégias que emergiriam no sentido de apoiá-la. De modo complementar a isso, a educação é, em sua própria constituição, uma área multidisciplinar, por isso alimentada por uma série de disciplinas que são, pode-se dizer, como suas artérias. Entre elas estão a filosofia, a psicologia, a antropologia, a sociologia, o estudo das mídias e das tecnologias e daí por diante. Conseqüentemente, se uma, ou algumas dessas áreas, vem passando por transformações em suas bases, isso pode afetar a própria fundamentação das teorias educacionais.

Acredito que realizar uma pesquisa situada nos pontos tangentes entre cognição, aprendizagem e tecnologias educacionais digitais se torna ainda mais oportuno quando, associada à ideia de que *máquinas podem aprender*, encontra-se uma outra hipótese, dialógica a essa, de que *máquinas podem ensinar*. Não considero tais questões meramente semânticas, mas também conceituais. Acredito, ainda, que se desdobram em crenças que acabam por embasar questões práticas. E essas são, então, em si, inquições relevantes quando se perspectiva que, na educação, aquilo que se entende por aprender e por ensinar não está dado; está em debate. Trata-se, justamente, de inquietações que impulsionam e também estabelecem limites e parâmetros para grande parte das pesquisas no campo. É a partir daí que investigo a concepção de aprendizagem que pode emergir principalmente a partir do enativismo. Procuro, assim, lançar luz crítica à concepção de aprendizagem presente em machine learning, que assume pressupostos parecidos com a concepção preponderante de aprendizagem na educação – o que traz oportunidades, mas também riscos, dos quais falarei.

Argumentação

A concepção de aprendizagem que emerge da perspectiva enativista para a cognição acentua elementos como a corporificação, o acoplamento do corpo com o mundo, a experiência e o fazer sentido do mundo individual e coletivo. A aprendizagem humana é enraizada na experiência, e essa experiência é concebida como o resultado da movimentação de um organismo autônomo no mundo ao qual

ele se acopla para sobreviver. O ser capaz de realizar cognição é concebido como um agente cognitivo autônomo, um *sense-maker*, que dá sentido aos espaços que ocupa justamente a partir da ocupação desses espaços. Por outro lado, assim o agente também modifica, molda esses espaços. O agente cognoscente é capaz de se engajar em relações de interação social com outros *sense-makers*, num processo chamado de *participatory sense-making*. Essa concepção contrasta com a concepção de aprendizagem presente em machine learning, que não compreende aquilo que se entende por *autonomia*, *sense-making* e experiência no enativismo. A aprendizagem de máquina se traduz em estatística, e seus processos automatizados baseados em dados deixam de fora elementos cruciais tanto para o funcionamento como para a compreensão das operações mentais humanas.

A partir da confrontação entre os pressupostos subjacentes à aprendizagem de máquina e aqueles que são subjacentes à aprendizagem humana segundo o enativismo, pode-se perceber potencialidades e limites da participação de sistemas de machine learning na aprendizagem humana. Um elemento me destaque nesta argumentação é o risco de redução da autonomia humana ao lidar com ML. Também se pode vislumbrar alguns caminhos para que essa parceria corra de maneira frutífera, buscando exemplos em que se consegue inserir as tecnologias na aprendizagem de maneira harmonizada com a necessidade de incluir o corpo, a afetividade e a experiência na aprendizagem. Há alguns cuidados que precisam ser tomados para que não se assuma uma equivalência entre determinados avanços cerebrais mensuráveis e a aprendizagem de fato, a qual envolve elementos que podem ser difíceis de ser mensurados; assim como não se pode assumir que a cognição se limita ao cérebro. A compreensão dos limites da aprendizagem de máquina e a compreensão da aprendizagem humana enraizada no enativismo compõem uma análise que nos apoia no entendimento de como sistemas de machine learning podem estar presentes na aprendizagem de modo que possamos avançar na inserção de tecnologias algorítmicas num contexto educacional que não pode prescindir do quanto já evoluiu no que diz respeito ao corpo, às emoções, ao ambiente cultural e social e às relações de troca entre cognoscentes.

A tese contém cinco capítulos, numerados de dois a seis. A seguir, faço uma breve apresentação dos capítulos.

No capítulo dois, intitulado ‘Máquinas para aprender, máquinas para ensinar – Uma perspectiva da aprendizagem automatizada no contexto da educação’, há

dois focos principais. O primeiro é uma análise do sentido intrínseco à aprendizagem em machine learning e o segundo volta-se para as primeiras máquinas de ensinar, usadas para automatizar o ensino de conteúdo aos alunos e testá-los em disciplinas. Qual a conexão entre esses focos e por que começar por eles? Primeiro, para investigar as bases daquilo que hoje se concebe como grande novidade na educação, mas que na verdade conta com uma estrutura e se baseia em premissas muito semelhantes às dessas máquinas de décadas atrás. As antigas máquinas de ensinar, que evoluíram para plataformas adaptativas, tinham como objetivos centrais atender aos alunos individualmente e deixar o professor livre de tarefas repetitivas para se concentrar em outras. Não são objetivos tão diferentes das plataformas desenvolvidas hoje que, potencializadas por inteligência artificial e *big data*, ganham novos contornos sem dúvida, mas continuam carregando as mesmas pressuposições subjacentes. Essas motivações, tão parecidas, podem revelar aspectos fundamentais da concepção de cognição que predominavam e talvez ainda predominem na educação.

Mesmo que a aprendizagem de máquina não seja equiparada à aprendizagem humana no campo de pesquisa, por vezes elas são comparadas. Um ponto crítico do discurso de que as máquinas aprendem como nós aprendemos é que, pela lógica, isso implicaria em que o contrário fosse igualmente verdadeiro, o que seria uma pressuposição problemática. Afinal, equiparando-se a aprendizagem humana à capacidade das tecnologias cognitivas, se poderia acabar concluindo que essas tecnologias são capazes não só de aprender, mas de ensinar. Uma vez que aprendizagem em ML se refere a fazer previsões com base em estatística, usando grandes bases de dados (*big data*), vêm sendo levantadas críticas por parte de estudiosos das tecnologias educacionais quanto a um possível condicionamento da aprendizagem a partir desses sistemas (*machine behaviourism*). Passo também por essas críticas, para explicitá-las e para reforçar que as de hoje acabam se parecendo com aquelas que foram direcionadas às máquinas analógicas. Essa história tem sido pouco contada, pois comumente se associa os primórdios da inteligência artificial na educação não a equipamentos movidos a manivelas e cartões, mas aos computadores em suas versões iniciais.

No decorrer da tese, procuro mostrar que a aprendizagem de máquina, baseada em estatística, deixa de fora aspectos fundamentais da cognição humana, tomando-se o enativismo como abordagem para a compreensão da mente. Por conta

disso, considero bastante pertinente esse mergulho inicial nas máquinas de aprender/de ensinar e as discussões sobre o que se concebe como aprendizagem nesse cenário. Trata-se, afinal, de um trabalho que deseja investigar pressupostos que sustentam a maneira como fazemos educação, hoje. Máquinas novas, potentes, mas criadas para trabalhar sobre as mesmas bases cognitivistas que nos sustentam e orientam como sociedade há décadas podem até contribuir para resolver os problemas que existem dentro desse contexto. Mas seu sucesso poderia ser posto em xeque se adotássemos outra perspectiva, a de um paradigma que compreende a cognição para além do cérebro e, especialmente, para além de um cérebro processador. E isto leva ao capítulo seguinte.

No capítulo três, intitulado ‘Tecnologias cerebrais’, o tema é cognição, neurociência e o paradigma ainda predominante para o estudo da mente humana, o cognitivista. A tecnologia digital fundamentada na ciência de dados tem o cérebro como premissa. A popularização da neurociência cresce em paralelo às novas abordagens à cognição. O que tem se modificado com relação à concepção da mente humana? Como isso pode afetar a maneira como concebemos essas tecnologias na educação? É este o problema investigado na seção. Argumento que as tecnologias de IA são desenvolvidas a partir do cérebro, negligenciando-se outros elementos que possivelmente constituem a mente humana. Por mais que a neurociência hoje já considere elementos externos ao cérebro como parte da cognição, ela ainda não atribui o mesmo peso a outros elementos que possivelmente constituem a mente, como o enativismo coloca. As tecnologias baseadas em machine learning pressupõem, por exemplo, que aprender é assimilar conteúdo; que a cognição é, simplesmente, atividade cerebral; que, compreendendo o cérebro, compreende-se como melhorar a cognição e, conseqüentemente, a aprendizagem; que aprender mais rápido é um ideal a ser buscado e contribui para que o processo de aprendizagem seja mais efetivo. Daí se conclui que o paradigma predominante na cognição ainda é o cognitivismo, e esse paradigma se estende à concepção de aprendizagem que ainda tem destaque hoje.

No decorrer ainda do capítulo três, introduzo ideias de Hubert Dreyfus, essenciais para a crítica filosófica à IA na aprendizagem que aqui procuro traçar. Acredito que a perspectiva de Dreyfus para a cognição pode ser vista como um impulso para a compreensão da mente humana a partir da proposta do enativismo, ainda que, na maior parte do tempo de atuação desse filósofo, ainda não estivesse

tão claramente definido esse arcabouço teórico reconhecido como os 4Es da cognição. Dreyfus é uma espécie de precursor das abordagens corporificadas, enativas, situadas à mente humana, e particularmente inspira este trabalho de pesquisa por ter como foco mais a inteligência humana do que a inteligência artificial, no fim das contas. Tendo Dreyfus reagido a um paradigma cognitivista/conexionista ao construir suas críticas à inteligência artificial, sinalizando que há mais na cognição humana do que regras, cálculos e representações, seu trabalho pode ser considerado ainda atual com relação aos desafios que a IA enfrenta – desafios que sempre correram paralelos aos que a ciência cognitiva tem enfrentado para compreender a mente humana. Também não estão distantes do que se deseja compreender sobre aprendizagem e sobre a participação do corpo e das emoções nesses processos.

Seguindo o percurso traçado até então, e após levantar algumas das questões centrais postas por Dreyfus que nos levam a buscar uma concepção diferente para a atividade cognitiva humana, no capítulo quatro, intitulado ‘A cognição enativa – A mente e(é) a vida’, introduzo as ideias e conceitos centrais do enativismo. São explicitados os conceitos de *sense-making*, *autonomia*, *emergência*, *experiência* e *corporificação*, e elaboradas as bases para que se possa reconhecer o contraste entre essa abordagem e a perspectiva cognitivista, representacionalista e computacional para a mente humana. Um dos conceitos, o de autonomia, é explicitado como fundamental para a ligação entre o enativismo e sua ideia chave – a mente como extensão da vida. Autonomia vem da ideia de autopoiese desenvolvida por Maturana e Varela (2019), acrescida, na versão enativista, do ingrediente da adaptividade. Com isso, é preparado o terreno para uma investigação que é apresentada na seção seguinte, sobre autonomia e machine learning.

O capítulo cinco, intitulado ‘Enativismo, autonomia e machine learning’, dedica-se a investigar o problema da autonomia do ser cognoscente ao lidar com sistemas de machine learning. Nos debates sobre a inteligência artificial e a sociedade, a discussão sobre os limites da autonomia de sistemas artificiais tem se mostrado vasta; mas, como fica a autonomia do ser cognoscente no contexto daquela que aqui chamo de *algoritmosfera*? Ele consegue se manter autônomo em meio a sistemas algorítmicos? E, se não consegue, quais as ameaças que sua autonomia sofre – e por quê? Partindo dessas perguntas impulsionadoras – e tendo em conta que neste momento a lógica das máquinas que aprendem alcança diversos

sistemas, numa rede que vai além das plataformas exclusivamente criadas para o ensino – argumento que, no acoplamento entre o cognoscente e o meio populado por algoritmos de ML, falta uma real interação. O encontro entre o cognoscente e o sistema artificial é desequilibrado, sob a perspectiva enativista da autonomia. Além disso, a circularidade dessas transações, nas quais o cognoscente fornece dados e é alimentado pela própria rede que nutre, em consequência provoca uma redução na variedade de experiências vividas pelo cognoscente. Isso reduz as chances de que o ser permaneça autônomo no máximo sentido possível, já que essa potencialidade depende da possibilidade de fazer escolhas. Além disso, considerando que nessas situações as normas não estão claras para o ser cognoscente, emerge um cenário obscuro para a autonomia do cognoscente, já que, para agir no mundo de uma maneira que garanta a manutenção da sua autonomia, é necessário conhecer as regras e os possíveis riscos a que se está exposto. Conclui-se que, quando seres cognoscentes usam sistemas de machine learning, sua autonomia (no sentido do enativismo) é ameaçada.

No capítulo seis, que é o último e se intitula ‘Abrindo a caixa preta’, em busca de uma perspectiva acerca de como se poderia contornar as ameaças à autonomia do cognoscente na algoritmosfera, proponho uma argumentação em duas etapas. A primeira traz uma série de premissas inerentes à concepção de aprendizagem que emerge a partir do enativismo, no intuito de encapsular essas premissas, que de certa forma aparecem no decorrer da tese. A segunda parte foca na corporificação e na intersubjetividade como elementos essenciais para que busquemos, na educação, orientações para a inserção de machine learning na aprendizagem a partir do enativismo.

2

Máquinas para aprender, máquinas para ensinar: uma perspectiva da aprendizagem automatizada no contexto da educação

Ensinar não é transferir conhecimentos, conteúdos (...) Não há docência sem discência, as duas se explicam e seus sujeitos, apesar das diferenças que os conotam, não se reduzem à condição de objeto, um do outro. Quem ensina aprende ao ensinar e quem aprende ensina ao aprender. (Paulo Freire)

Machine learning, ou aprendizagem de máquina, consiste em uma vertente da inteligência artificial que vem se entranhando nos domínios do nosso cotidiano, muitas vezes sem que nos demos conta de sua presença. O universo da educação não fica de fora. Cada vez mais presente em sistemas que já são ou que virão a ser utilizados na educação, presencial ou online/a distância, a aprendizagem de máquina é encontrada, por exemplo, na forma de *chatbots*, nas plataformas adaptativas de ensino e aprendizagem e em sistemas criados para automatizar avaliações. Em nosso dia a dia, conectados à internet, estamos ligados a sistemas de ML, alimentados pelos nossos próprios dados. A aprendizagem de máquina está por trás das redes sociais em que navegamos, dos sites que acessamos, dos aplicativos para celular e dos sistemas digitais de bancos, instituições, sistemas de saúde, entre muitos outros. Trata-se de soluções pensadas para fins específicos – isto é, em vez de esses sistemas tentarem dar conta da inteligência humana como um todo, o que tendia a ser o foco das pesquisas iniciais em IA, a aprendizagem de máquina volta-se para determinadas categorias de tarefas. Para endereçá-las, são desenvolvidos os algoritmos; para que os algoritmos possam operar e se manter em movimento, sistemas de ML processam *big data*: grandes massas de dados gerados pela atividade dos usuários.

Diz-se, então, que a aprendizagem de máquina se utiliza de dados para “aprender”. Mas qual é, de fato, o sentido de “aprendizagem” nesse contexto? Qual o conceito de aprendizagem que subjaz a esses sistemas? O que se quer dizer, de fato, quando se faz referência a uma máquina que aprende? A aprendizagem de máquina se dá da mesma maneira que a aprendizagem humana? Um sistema de ML poderia, de fato, proporcionar aprendizagem a um ser humano, isto é, levar um ser humano a aprender? Antes de buscar respostas a tantas perguntas, é preciso olhar para o sentido de aprendizagem vinculado a esses sistemas de inteligência artificial

(IA)⁹ e então colocá-lo em perspectiva. Neste capítulo, procuro encaminhar essa compreensão, a qual estrutura em uma argumentação que vai se complementar com o capítulo seguinte.

Divido este capítulo em duas partes. Na primeira, procuro explorar a razão pela qual sistemas de machine learning se chamam assim. Apresento também alguns dos usos ou aplicações de sistemas algorítmicos na educação. Na segunda parte, estabeleço paralelos entre as máquinas que aprendem, de hoje, e as máquinas criadas para ensinar, há muitas décadas. Faço isso para explicitar semelhanças entre as justificativas que impulsionam o uso de plataformas baseadas em ML na educação hoje e as motivações que delineavam antigos recursos para a automação da educação – as máquinas analógicas. Essas motivações, tão parecidas, podem revelar aspectos fundamentais da concepção de aprendizagem que predominava e tende a ainda predominar na educação.

Ao final do capítulo, trago algumas das questões críticas pertinentes à aprendizagem de máquina na educação a partir do olhar de pesquisadores em tecnologias educacionais.

2.1

Breve definição de machine learning

Machine learning ou aprendizagem de máquina por vezes é identificada como sinônimo de inteligência artificial. No entanto, mais precisamente, essa é uma subárea da IA (Holmes, Bialik & Fadel, 2010; Nilsson, 2010) que não surgiu de repente e nem é tão recente quanto se possa ser levado a crer. Segundo Mackenzie (2015), o reconhecimento de padrões, a modelagem estatística, a aprendizagem de máquina já eram campos ativos de pesquisa desde antes da II Guerra Mundial, ainda que se limitassem a certas áreas de aplicação específicas da ciência, da administração pública e da indústria, além de algumas áreas comerciais como a avaliação do risco de crédito nos anos de 1960.

As técnicas de aprendizado de máquina quase todas giram em torno de maneiras de transformar, construir ou impor algum tipo de forma aos dados e usar essa forma

⁹ Para simplificar, por vezes refiro-me a inteligência artificial como IA.

para descobrir, decidir, classificar, agrupar, recomendar, rotular ou prever o que está acontecendo ou o que acontecerá. (Mackenzie, 2015, p. 432)¹⁰.

Segundo Mitchell (2021), foi nos anos de 1900 e 2000 que a aprendizagem de máquina ou machine learning – “o desenvolvimento de algoritmos que criam modelos preditivos a partir de dados” – deu um salto “meteórico. “Essas abordagens eram inspiradas por estatísticos, mais do que pela neurociência ou pela psicologia, e eram voltadas para o desempenho de tarefas específicas em vez de dar conta da inteligência generalista”¹¹. Por volta de 2010, emergiu com força total a subárea da aprendizagem profunda ou *deep learning*, “em que redes neurais multicamadas, inspiradas no cérebro humano, são treinadas a partir de dados”¹² (Mitchell, 2021, p. 2).

Ao pensar em computadores, pensamos em programação. Pensamos em um passo a passo que o computador deveria seguir para cumprir determinada tarefa, contendo regras para ser seguidas à risca. De fato, assim é a programação em sua forma “clássica” e era a IA em sua primeira fase. Sistemas de ML, por sua vez, não demandam que lhes sejam dadas todas as direções previamente. Ainda que os algoritmos sirvam para dar “instruções” a esses sistemas (Domingos, 2015, p. 1), são instruções sensivelmente diferentes. “Em vez de algoritmos serem programados quanto ao que exatamente deverão fazer, de modo geral eles têm a habilidade de *aprender* o que fazer” (Holmes, Bialik & Fadel, 2010, p. 89, grifo meu). Os autores esclarecem que isso, porém, não significa que ML não exija grandes esforços de programação. O que acontece é que o empenho vai em uma direção diferente: em vez de se desenvolverem comandos diretos que levem a *outputs* (resultados) diretos, programação em ML envolve levar grandes massas de dados (*inputs*) a gerar *outputs* a partir de previsões.

Sendo assim, sistemas de ML são capazes de construir hipóteses a partir de dados. Fazem inferências “indutivas”, em vez de “dedutivas”. Por exemplo

¹⁰ No original: *The techniques of machine learning nearly all pivot around ways of transforming, constructing or imposing some kind of shape on the data and using that shape to discover, decide, classify, rank, cluster, recommend, label or predict what is happening or what will happen.* (Mackenzie, 2015, p. 432)

¹¹ No original: *The 1990s and 2000s saw the meteoric rise of machine learning: the development of algorithms that create predictive models from data. These approaches were typically inspired by statistics rather than by neuroscience or psychology, and were aimed at performing specific tasks rather than capturing general intelligence.* (Mitchell, 2021, p. 2)

¹² No original: *brain-inspired multilayered neural networks are trained from data.* (Mitchell, 2021, p. 2)

(Nilsson, 2010, p. 398), se um grande conjunto de dados contém vários casos de cisnes brancos, e nenhum caso de cisnes de outras cores, um algoritmo pode fazer a inferência de que “todos os cisnes são brancos”. Trata-se de uma inferência indutiva porque, enquanto as inferências dedutivas seguem premissas necessariamente e de forma lógica, as indutivas são hipóteses, “sempre sujeitas à falsificação ou a dados adicionais”. Pode haver, quem sabe, uma ilha ainda desconhecida somente com cisnes negros que, se descoberta, irá contradizer a inferência. Mesmo assim, há uma utilidade em inferências indutivas baseadas em grandes conjuntos de dados. “A ciência em si se baseia em inferências indutivas” (Nilsson, 2010, p. 398). Sistemas de ML sempre funcionam a partir de grandes massas de dados (*big data*). Ainda segundo Nilsson (2010), não é incomum esses sistemas classificarem previamente, de alguma maneira, os dados com os quais operam; porém, na medida em que a memória dos computadores consegue acessar informações de maneira cada vez mais rápida, menos técnicas de classificação *a priori* são necessárias. Uma redução, ou agrupamento, pode acontecer no próprio momento em que uma determinada decisão precisa ser tomada.

“Algoritmos de machine learning analisam dados para identificar padrões e construir um modelo que é então usado para prever valores futuros” (Holmes, Bialik & Fadel, 2019, p. 89). Para ilustrar, os autores mencionam a identificação de padrões históricos em dados do mercado de ações e a previsão de movimentações futuras; padrões em fotografias de pessoas nomeadas, prevendo quem é mostrado em outras fotografias; e padrões em sintomas relacionados à saúde, prevendo um diagnóstico específico. Algoritmos são treinados para reconhecer vozes, rostos e imagens; fazer recomendações; dirigir carros de forma autônoma, entre outras finalidades (Knox, Williamson & Bayne, 2019). São recursos presentes em softwares de filtragem de SPAM em nossos serviços de e-mail; em sistemas de recomendação de filmes, livros e outros produtos em sites de compra e plataformas de *streaming*; em aplicativos de ensino de idiomas (alguns desses sistemas são especificados e nomeados em Holmes, Bialik & Fadel, 2019); em robôs capazes de circular pelos andares de um hotel ou de uma fábrica; em dispositivos capazes de “conversar” com um ser humano. Segundo Domingos (2015), não é exagero dizer que os algoritmos formam a base do mundo que conhecemos hoje:

Eles estão no tecido da vida cotidiana. Não estão apenas em seu telefone celular ou no seu laptop, mas no seu carro, na sua casa, nos seus dispositivos, nos seus

brinquedos. Seu banco é um emaranhado gigante de algoritmos, com humanos apertando botões aqui e ali. Algoritmos programam voos e depois fazem os aviões voarem. Algoritmos comandam fábricas, comércio e rotas de mercadorias, descontam rendimentos, armazenam registros. Se todo algoritmo de repente parasse de funcionar, seria o fim do mundo como o conhecemos. (Domingos, 2015, p. 1)¹³

Haveria muitas especificações técnicas a esmiuçar acerca da aprendizagem de máquina; porém, para os fins do recorte aqui desejado, a seguir apresento uma explicação breve e simplificada dos três tipos mais comuns, a partir de Holmes, Bialik e Fadel (2019, pp. 90-93) e de Knox, Williamson e Bayne (2019), e em seguida explico brevemente o conceito de *big data*.

Aprendizagem Supervisionada

Na chamada aprendizagem supervisionada, o sistema recebe grandes quantidades de dados para os quais os *outputs*, ou resultados, já são conhecidos ou esperados. Isto é, os dados já foram previamente classificados para a construção do modelo preditivo. O que o sistema precisa fazer é relacionar dados que recebe com rótulos que já “conhece”. Essa é a abordagem usada para identificar pessoas em fotografias que os usuários sobem para seus perfis em redes sociais (a máquina usa milhões de fotos submetidas pelos usuários e rotuladas por eles mesmos para identificar e rotular automaticamente essas mesmas pessoas). É também esse tipo de aprendizagem de máquina que está presente naquelas imagens que frequentemente somos solicitados a identificar, a partir de um comando específico (Holmes, Bialik & Fadel, 2019, p. 90) – por exemplo, *marque todas as imagens que têm semáforos, aviões, ou placas de rua* – quando estamos prestes a entrar num site que deseja saber se somos “humanos” (curiosamente, neste momento fazemos uma tarefa porque somos humanos justamente para que máquinas possam fazê-las também!).

¹³ No original: *They are woven into the fabric of everyday life. They're not just in your cell phone or your laptop but in your car, your house, your appliances, and your toys. Your bank is a gigantic tangle of algorithms, with humans turning the knobs here and there. Algorithms schedule flights and then fly the airplanes. Algorithms run factories, trade and route goods, cash the proceeds, and keep records. If every algorithm suddenly stopped working, it would be the end of the world as we know it.* (Domingos, 2015, p. 1)

Aprendizagem Não Supervisionada

Na aprendizagem não supervisionada, o sistema procura identificar nos dados padrões marcantes, bem característicos, para então desenvolver uma análise a partir do “caos”. Isto é, o sistema recebe uma carga (ainda maior) de dados, mas desta vez eles não foram classificados previamente: não estão rotulados. Por meio da análise desses dados, os algoritmos “descobrem” padrões escondidos na estrutura sobre a qual os dados se situam, ou definem grupos de dados que podem ser usados para classificar novos dados. Assim, por exemplo, o Google reconhece pessoas em fotografias. Mas há outros casos, como a divisão de consumidores em grupos para que recebam anúncios de acordo com suas preferências; a identificação de letras e números a partir da exposição a manuscritos e o treinamento de sistemas para distinguir uma transação fraudulenta de uma legítima (exemplos em Holmes, Bialik & Fadel, 2019, p. 90-91).

Reinforcement Learning ou Aprendizagem por Reforço

Neste tipo de aprendizagem de máquina, o sistema recebe dados iniciais a partir dos quais gera seu modelo; esse modelo é classificado como correto ou incorreto, e então se recompensa ou se pune de acordo. A IA usa esse reforço negativo ou positivo para atualizar seu modelo e tentar novamente, e assim desenvolve-se – “aprendendo e evoluindo” com o tempo (Holmes, Bialik & Fadel, 2019, p. 91). Um exemplo que os autores dão é o do carro automático que evita uma colisão; o modelo que o permitiu evitar bater é reforçado e recompensado, de modo que isso aumenta a “habilidade” do sistema de evitar colisões no futuro. Segundo Knox, Williamson e Bayne (2019), este tipo de aprendizagem de máquina é uma das mais proeminentes áreas de ML hoje, e é a técnica usada no jogo AlphaGo, da Deep Mind¹⁴.

2.1.1 Big Data

Como já explicitado, dizemos que sistemas de ML exigem um tipo específico de programação porque, diferente dos sistemas que precisam de códigos como um

¹⁴ Ver <https://deepmind.com/research/case-studies/alphago-the-story-so-far>; acesso em 5 de junho de 2021. Há também um documentário sobre o AlphaGo neste link: <https://www.youtube.com/watch?v=WXuK6gekU1Y>

passo a passo para pautar suas ações do início ao fim, eles funcionam a partir de parâmetros: uma programação que é como um pontapé para o sistema funcionar – mas que depende dos dados obtidos a partir dos usos desses sistemas para seguir operando. Sendo assim, não há como falar em machine learning sem falar em *big data*: os conjuntos maciços de dados de usuários (Ferreira et al., 2018) que abrem o caminho para que as redes neurais profundas possam operar.

Big data virou um termo polêmico que grupos diferentes utilizam em diversos sentidos. A definição técnica simples é que os big data oferecem uma enorme quantidade de informação, extremamente diversificada, reunida a uma velocidade extraordinária. Em vez de trabalhar com uma definição técnica restrita de big data, ao fim se pode compreendê-lo melhor como um fenômeno social emergente e como um conceito de enorme força que tem adquirido importância enorme nos últimos anos. Os macrodados são também inseparáveis dos softwares, dos algoritmos e da análise necessária para os reunir e gerenciar, o que demanda vários tipos de especialistas. (Williamson, 2017, p. 15)

Mitchell (2021, p. 2-3) postula que, apesar de as redes neurais profundas existirem desde a década de 1970, foi somente graças à possibilidade de utilizar vastas bases de dados que esses tipos de sistemas se tornaram populares – daí a inseparabilidade entre big data e algoritmos a que Williamson se refere na citação acima. Ainda segundo Mitchell, outros atributos essenciais para as redes neurais profundas decolarem foram o desenvolvimento de chips de processamento paralelo rápido e as inovações nos métodos de treinamento das máquinas.

Uma vez que os sistemas de machine learning operam a partir de dados de seus usuários, trata-se de um processo retroalimentado: os usuários fornecem os dados, os sistemas consomem esses dados e geram novos conteúdos que gerarão novos dados, e assim por diante. Pode-se dizer, então, que a expansão da capacidade das redes neurais, de um modo geral, está diretamente ligada à multiplicação das redes sociais online e à popularização dos smartphones e de outros dispositivos digitais, que tornam cada vez mais fácil e rápido o compartilhamento de informações online.

Dado o crescimento exponencial do poder dos computadores, a disponibilidade de dados (big data) devido às mídias sociais e o uso massivo de bilhões de smartphones, e as redes móveis ágeis, a IA, especialmente machine learning, tem feito progresso significativo (Coeckelberg, 2020, p. 3)¹⁵

¹⁵ No original: *Given the exponential growth of computer power, the availability of (big) data due to social media and the massive use of billions of smartphones, and fast mobile networks, AI, especially machine learning, has made significant progress.* (Coeckelberg, 2020, p. 3)

Como Véliz (2020)¹⁶ indica, a “economia de dados” gira a partir dos aparentemente mais ínfimos atos que protagonizamos cotidianamente, e se constitui a partir da perda progressiva da nossa privacidade. “Qual a primeira coisa que você faz quando acorda? Você provavelmente checka seu telefone. *Voilà*: pela primeira vez no seu dia, você perde pontos no que diz respeito aos seus dados” (Véliz, 2020, p. 8). Usando o smartphone logo pela manhã, informamos a uma série de companhias interessadas em nossos dados a que horas levantamos, onde dormimos ou com quem. Para ilustrar como a coleta de dados segue no decorrer do dia e da noite, em casa ou em locais públicos, a pesquisadora aponta exemplos como o do *smart watch*, que passa a noite enviando dados como a frequência cardíaca de quem o carrega no pulso; o das câmeras de reconhecimento facial nas ruas e o da *smart TV* que, ligada a smartphones, ao ser acionada permite por exemplo que um parceiro perceba que o outro ainda não saiu de casa para o trabalho.

A popularização da prática de compartilhar conteúdo online provoca um movimento espontâneo de geração de dados, no sentido de que os usuários fazem isso sem nem mesmo se darem conta, e o fazem o tempo todo. No universo online, misturam-se o uso dos sistemas e o compartilhamento de informações que se tornam dados para esses sistemas. Mas pode-se também obter dados a partir de ambientes, digamos, “controlados”, como sistemas de machine learning utilizados por grupos específicos e para fins específicos. Um exemplo ilustrativo é a IRIS+, instalada no Museu do Amanhã do Rio de Janeiro depois de a mesma tecnologia ser usada em museus de São Paulo. Por trás do sistema está o recurso da IBM chamado de Watson. Ao final da visita às exposições no local, os transeuntes podem acessar o sistema e “conversar” com a IA, que “questiona” o que foi que mais preocupou o visitante durante o passeio (o museu tem a característica de problematizar questões climáticas, sociais, culturais etc). A partir da resposta é fornecida então ao visitante uma lista de opções de engajamento que ele poderia ter com instituições que trabalham para resolver aqueles problemas que mais o

¹⁶ Carissa Véliz é pesquisadora e professora da Universidade de Oxford. Em seu livro “Privacy is Power – Why and How You Should Take Back Control of your Data” (2021), em que procura orientar os leitores quanto à proteção de seus dados, ela usa o termo “data economy” para se referir à coleta de dados feita em modo contínuo por governos e corporações; o termo, segundo ela, é intercambiável com outro, “surveillance capitalism” ou capitalismo de vigilância (ver também Zuboff, 2020). Vale a pena consultar a parte final desse livro de Véliz para conhecer uma lista de orientações que a autora elabora no sentido da proteção de dados pessoais na era do *big data*.

preocuparam. Para dar início ao uso da tecnologia de IA no museu, uma equipe cadastrou diversas possibilidades de respostas e dicas correspondentes; com o tempo, as respostas dos usuários passaram a ser incorporadas. Apesar de a IA ter a capacidade de buscar informações em grandes bases de dados online, ela não dispensa um processo de curadoria, já que o museu avalia ter responsabilidade sobre o conteúdo transmitido aos visitantes¹⁷.

2.2 Inteligência artificial na educação

Há indicações de que a história da inteligência artificial na educação formal começa na década de 1980, tendo como marco a publicação da primeira edição do *International Journal of Artificial Intelligence in Education* em 1989 (na qual apareceu o trecho que citei na introdução da tese, em que Schank e Edelson destacam as questões que atam fundamentalmente a pesquisa em IA à pesquisa em educação) e a criação da International AI in Education Society (IAIED) em 1993, como pontuam Williamson e Eynon (2020, p. 224). Os sistemas inteligentes de tutoria e instrução assistida por computador (CAI, na sigla em inglês para Computer-Assisted Instruction) teriam sido, segundo esses autores, os antecessores da IA na educação, a qual vem se desdobrando em duas vertentes básicas de aplicação: o desenvolvimento de ferramentas a serem utilizadas em sala de aula e o uso de IA para medir, compreender e melhorar o ensino, como destacam Holmes, Bialik e Fadel (2019). Segundo Pea (2016), nos anos de 1980 houve uma intensa mobilização de diversos grupos em torno dos possíveis usos da informática na educação. Isso teria a ver com o campo da ciência cognitiva, então ainda incipiente, e com a crescente popularização do acesso às tecnologias. O clima era de otimismo.

Durante esta era, houve uma série de reuniões diferentes nos encontros da American Educational Research Association (AERA), da Association for the Advancement of Artificial Intelligence e da Cognitive Sciences Society, e até mesmo nas conferências do Computer Human Interaction e Computer Supported Cooperative Work, onde pesquisadores em universidades, laboratórios e na indústria começaram a falar sobre como reunir as forças que tomam forma nessas interseções férteis entre cognição,

¹⁷ Com informações fornecidas pela equipe do museu para que seu escrevesse o texto disponível em <https://porvir.org/inteligencia-artificial-museus-querem-se-aproximar-do-visitante/>

computação e educação, e estavam compartilhando suas últimas descobertas de pesquisa e demonstrações de suas tecnologias. (Pea, 2016, p. 38)¹⁸

Segundo Schank e Edelson (citados em Williamson & Eynon, 2020), a IA e as pessoas por trás desses sistemas teriam um impacto positivo na educação, tanto na questão de organização curricular e de design instrucional como nas maneiras como o ensino e a aprendizagem seriam compreendidos e praticados. Outros depois deles defenderam os aspectos positivos de aumentar a velocidade e a qualidade do aprendizado com o uso de IA, o que recrutaria as empresas de tecnologia; o ensino individualizado e a aprendizagem invertida ou por projetos sempre estiveram em pauta nessas defesas, como fatores supostamente a transformar a escola e oferecer novos objetivos para as práticas de ensino e aprendizagem (McArthur, Lewis & Bishay, 1995, apud Williamson & Eynon, 2020).

Na educação, formal ou informal, a inteligência artificial vem aparecendo em diferentes formatos. Um deles é o da aprendizagem de máquina na forma de plataformas de aprendizagem adaptativa – que interessam particularmente a este capítulo, por dialogarem com as primeiras máquinas de ensinar. Mas há outros tipos de recursos, como agentes inteligentes em aprendizagem baseada em jogos; *chatbots*; aplicativos em celulares e a tutoria online no acompanhamento de tarefas. Como Williamson (2017) explicita, são quatro os objetivos da educação baseada em dados, segundo o Center For Data Innovation¹⁹: a personalização; a aprendizagem baseada em evidência; a eficiência da escola e a inovação contínua. “A personalização talvez tenha se convertido na principal palavra-chave desse tipo de educação, com ênfase em sistemas e processos que podem se adaptar de forma inteligente a cada aluno” (Williamson, 2017, p. 30-31).

Os trechos a seguir foram extraídos de blogs de empresas de tecnologias educacional, sites voltados para a disseminação dos usos de tecnologias digitais na educação ou blogs relacionados a tecnologias educacionais em geral:

Com base no desempenho anterior, as plataformas de treinamento podem fornecer uma inteligência de dados perspicaz para prever as necessidades dos alunos, criar

¹⁸ No original: *During this era, there were a number of different convenings at meetings of the American Educational Research Association (AERA), the Association for the Advancement of Artificial Intelligence, and the Cognitive Sciences Society, and even the Computer Human Interaction and the Computer Supported Cooperative Work conferences, where researchers in universities, labs, and industry began talking about how to gather together the forces taking shape in these fertile intersections of cognition, computing, and education, and were sharing their latest research findings and technology demos. (Pea, 2016, p. 38)*

¹⁹ <https://datainnovation.org/about/> (Acesso em 10 de junho de 2021)

*experiências personalizadas e centradas no aluno e alcançar melhores resultados de aprendizagem (...) Melhoria da eficiência do curso através da análise preditiva - A principal vantagem da aprendizagem de máquina na educação é sua capacidade de acompanhar o progresso do aluno e ajustar os cursos para responder às necessidades reais dos alunos, aumentando assim o engajamento e oferecendo treinamento de alta qualidade. O feedback dos algoritmos de ML permite aos instrutores compreender o potencial e os interesses dos alunos, identificar alunos em dificuldades, detectar lacunas nas habilidades e fornecer apoio extra para ajudar os alunos a superar os desafios de aprendizagem.*²⁰

*A aprendizagem de máquina (ML) está transformando a educação e mudando fundamentalmente o ensino, a aprendizagem e a pesquisa. Os educadores estão usando ML para detectar estudantes em dificuldades mais cedo e tomar medidas para melhorar o sucesso e a retenção. Os pesquisadores estão acelerando a pesquisa com o ML para desbloquear novas descobertas e insights. ML está expandindo o alcance e o impacto do conteúdo de aprendizagem on-line através da localização, transcrição de texto para fala e personalização*²¹.

*O aprendizado de máquina apresenta definitivamente uma série de características vantajosas para os estudantes e seus professores. A possibilidade de uma educação personalizada que pode poupar tempo para os professores e ajudar a prever o sucesso futuro de um estudante é inestimável. No entanto, ainda temos um longo caminho a percorrer. Os educadores não vão querer comprometer as habilidades sociais de seus alunos, e devem considerar se a administração escolar deles pode arcar com o equipamento. As escolas terão que decidir como equilibrar as características que mais significam para eles e como eles podem se dar ao luxo de implementá-las*²².

A aprendizagem de máquina é o futuro do eLearning, pois pode oferecer aos alunos online uma variedade de benefícios. Vamos dar uma olhada nos benefícios que pode oferecer ao futuro do eLearning (...) 1. Melhora o Retorno do Investimento em eLearning (...) Com a ajuda de análises preditivas, você pode acompanhar o progresso de seus alunos e outras experiências de aprendizagem (...) 2. Oferece conteúdo de eLearning mais personalizado - Um dos benefícios da aprendizagem de

²⁰ <https://www.intellias.com/benefits-of-machine-learning-in-education/> (Acesso em 2 de junho de 2021); no original: *Based on past performance, training platforms can deliver insightful data intelligence to predict students' needs, create customized, learner-centered experiences, and achieve better learning outcomes (...) Improved course efficiency via predictive analytics - The major boon of machine learning in education is its ability to track learner progress and adjust courses to respond to students' actual needs, thus increasing engagement and delivering high quality training. Feedback from ML algorithms allows instructors to understand learners' potential and interests, identify struggling students, spot skill gaps, and provide extra support to help students overcome learning challenges.*

²¹ <https://aws.amazon.com/pt/education/ml-in-education/> (Acesso em 2 de junho de 2021); no original: *Machine learning (ML) is transforming education and fundamentally changing teaching, learning, and research. Educators are using ML to spot struggling students earlier and take action to improve success and retention. Researchers are accelerating research with ML to unlock new discoveries and insights. ML is expanding the reach and impact of online learning content through localization, transcription, text-to-speech, and personalization.*

²² <https://www.thetechadvocate.org/the-benefits-and-limitations-of-machine-learning-in-education/> (Acesso em 2 de junho de 2021); no original: *Machine learning definitely presents a number of advantageous features for students and their teachers. The possibility of a personalized education that can save time for teachers and help to predict a student's future success is invaluable. However, we still have a long way to go. Educators won't want to jeopardize the social skills of their students, and they must consider if their school district can afford the equipment. Schools will have to decide how to balance out the features that mean the most to them and how they can afford to implement them.*

máquinas é sua capacidade de usar o reconhecimento de padrões para prever os resultados da aprendizagem. Por exemplo, um algoritmo de aprendizagem de máquina identificará onde um aluno está lutando com seu módulo de curso. Após identificá-lo, ele automaticamente faz ajustes, fornecendo informações úteis para o progresso. Se um aluno apresentar uma lacuna de habilidade específica, o sistema automaticamente recomenda conteúdo em um formato personalizado para ajudar o aluno a adquirir conhecimento (...) 3. Utiliza Chatbots Como Instrutores de eLearning - Os Chatbots são softwares construídos com aprendizagem de máquinas e IA para interagir com humanos. Este software pode tomar o lugar de um instrutor de eLearning para proporcionar uma melhor experiência de aprendizagem (...) Os Chatbots são essenciais no eLearning, pois não têm restrições para responder perguntas ou passar projetos aos alunos e avaliá-los (...) 5. Fornece múltiplos formatos de avaliação (...) Com a introdução do aprendizado por máquina, você pode acessar o conhecimento de seus alunos com diferentes formatos de questionários²³.

Individualização do ensino, redução do trabalho “mecânico” do professor (como na correção e produção de exercícios e avaliações), respostas mais rápidas aos alunos, automação e aumento da eficiência no ensino e na aprendizagem: como se vê pelos trechos destacados, estas são algumas das principais motivações associadas ao uso de ML na educação, hoje. Curiosamente, essas motivações se assemelham a aquelas que acompanharam invenções muito anteriores aos computadores e à IA na educação. Consequentemente, os antecessores das tecnologias educacionais digitais baseadas em IA tais como as concebemos hoje podem estar localizados não nos computadores digitais, mas num ponto significativamente mais remoto da história das tecnologias do que se imagina: como indica Watters (2021), a origem de tudo isso estaria ligada às tentativas de automação do ensino que vêm se desenhando desde a década de 1920, pelo menos. Na seção a seguir, procuro trazer um pouco desse histórico, estabelecendo conexões

²³ <https://elearningindustry.com/machine-learning-benefits-elearning> (Acesso em 2 de junho de 2021); no original: *Machine learning is the future of eLearning as it can offer online learners a variety of benefits. Let's have a look at the benefits it can offer to the future of eLearning (...) 1. Improves eLearning Return On Investment (...) With help from predictive analytics, you can track your learners' progress and other learning experiences (...) 2. Delivers More Personalized eLearning Content - One of the benefits of machine learning is its ability to use pattern recognition to predict learning outcomes. For instance, a machine learning algorithm will identify where a learner is struggling with your course module. After identifying it, it automatically makes adjustments by providing helpful information to progress. If a student exhibits a specific skill gap, the system automatically recommends content in a personalized format to help the learner acquire knowledge (...) 3. Employs Chatbots As eLearning Instructors - Chatbots are software built with machine learning and AI to interact with humans. This software can take the place of an eLearning instructor to deliver a better Learning Experience (...) Chatbots are essential in eLearning as they have no restrictions on answering questions or giving projects/assessments to students (...) 5. Provides Multiple Assessment Formats (...) With the introduction of machine learning, you can access your learners' knowledge with different quiz formats.*

que considero ricas para compreendermos o momento que estamos vivendo e pensar criticamente os possíveis caminhos ainda por vir.

2.3 Programando o ensino

No livro “Ensino programado” – uma nova tecnologia didática” (1969), lançado a partir de sua tese de doutorado, a educadora e pesquisadora Vera Candau aborda o momento em que se teria começado a falar sobre a necessidade de individualizar o ensino e, mais especificamente, a didática para que houvesse uma “adequação metodológica ao psiquismo de cada indivíduo”. Ela se refere ao trabalho de Rousseau como um marco nesse sentido; porém registra que:

Já na antiguidade alguns educadores se preocuparam com adaptar seu trabalho docente às exigências do indivíduo enquanto tal. Mas é bem verdade que só na época moderna encontramos colocado com plena consciência o problema da adequação metodológica ao psiquismo de cada indivíduo. (Candau, 1969, p. 18-19).

Com relação às adaptações que a escola teria feito para cumprir com a necessidade de dar conta dos indivíduos e suas diferenças, Candau assinala entre as primeiras iniciativas o agrupamento em classes de alunos de capacidade aproximada. Como a autora aponta, há na educação um extenso histórico de investidas no sentido de endereçar diferentes “estados mentais” ou “idades mentais” dos estudantes. Não cabe aqui um aprofundamento dessas tentativas; o que considero importante pontuar, para os fins de analisar como as primeiras máquinas de ensinar se relacionam com as plataformas atuais de aprendizagem, é o fato de que a orientação para o ensino individualizado não é algo recente. Também não é nova a tentativa de utilizar máquinas para a aplicação de testes e atividades que permitam ao aluno uma dinâmica de pergunta-e-resposta instantânea, estratégia muito parecida com as de algumas plataformas baseadas em machine learning vigentes hoje²⁴. Nesse histórico, destacam-se os nomes de Sidney Pressey e B. F. Skinner.

Em 1926, Pressey, professor de psicologia educacional, desenvolveu um dispositivo mecânico para oferecer perguntas e respostas como um “meio mais

²⁴ Destaca-se ainda o imediatismo como ponto fundamental para o professor; obter conhecimento sobre o que a classe “aprendeu ou não” sobre determinado tema a partir da aplicação de testes com imediato retorno é apontado como vantagem de diversas plataformas educacionais baseadas em IA.

econômico e eficiente para a aplicação e correção de testes” (Candau, 1969, p. 31-32). A invenção do professor, que lecionava na Ohio State University, teria surgido em meio à ebulição – testemunhada então nos Estados Unidos – do desenvolvimento dos testes objetivos de inteligência e de instrução. Segundo Watters (2015), o próprio Pressey e sua esposa foram proprietários de diversos desses testes padronizados, nos anos de 1920. Para Pressey, o reforço imediato era essencial; ele acreditava que, para que houvesse aprendizagem, a resposta correta deveria vir o quanto antes para o aluno. “Dispositivos ou materiais especiais que, ao mesmo tempo, informam a um estudante sobre sua resposta a uma questão estar correta ou não e então o levam à resposta correta claramente fazem mais do que testá-lo; *eles também o ensinam*” (Pressey, 1950, p. 418, destaque original do autor)²⁵.

Como se vê no trecho destacado pelo próprio Pressey, nota-se que, em certa medida, ele não hesitava em atribuir a tais sistemas a capacidade de ensinar. Em reforço a esse argumento, e ressaltando as vantagens dos testes de resposta rápida sobre a própria capacidade do professor, Pressey discorreu sobre os benefícios de utilizar recursos dessa natureza junto aos estudantes. Entre eles estariam a grande quantidade de questões a serem disponibilizadas em uma mesma ocasião; a maior quantidade de temas cobertos; as instruções individualizadas possivelmente dirigidas a cada estudante seguidas de correções também individualizadas:

Um teste objetivo bem feito tem uma resposta correta mais cuidadosamente ponderada e mais precisamente formulada do que um professor provavelmente dará numa discussão em classe; e as respostas erradas são cuidadosamente escolhidas como os equívocos ou erros mais comuns (...) estes dispositivos devem permitir cobrir mais conteúdo em um determinado tempo, e de forma mais adequada, do que um professor poderia. (Pressey, 1950, p. 419)²⁶

Pressey criou diversos dispositivos, ou várias versões de uma mesma ideia básica, como relatam Holmes, Bialik e Fadel (2019); dentre eles, um que foi batizado de “*punchboard*” – é interessante ressaltar que foi a partir dos dados

²⁵ No original: *Devices or special materials which at once inform a student about the correctness of his answer to a question, and then lead him to the right answer clearly do more than test him; they also teach him.* (Pressey, 1950, p. 418)

²⁶ No original: *A well made objective test has a more carefully considered and exactly phrased right answer than an instructor is likely to give in class discussion; and the wrong answers are carefully chosen as the most common misconceptions or mistakes (...) these devices should make it possible to cover more ground in a given time than a teacher could and do it more adequately.* (Pressey, 1950, p. 419)

obtidos com essa invenção que Pressey redigiu seu artigo de 1950 citado aqui (o que mostra que, com aquilo que observava a partir de suas próprias invenções, ele tirava conclusões e gerava dados para seus trabalhos futuros). Candau (1969) oferece uma descrição do funcionamento do dispositivo, um instrumento a permitir “exercícios de perfuração”:

No punchboard o aluno precisa inserir o lápis em um dos quatro furos, conforme a resposta que haja escolhido como certa. Se tiver escolhido a que for certa, a matriz que fica por baixo, permitindo a entrada da ponta do lápis mais profundamente, registrará o acerto. Assim, o aluno ficará sabendo se sua resposta for exata, e o professor poderá ler na folha de respostas o número das respostas certas e erradas. As placas furadas e as matrizes de respostas podem ser substituídas por folhas de respostas quimicamente preparadas, onde a resposta certa, tocada com um lápis especial, mudará de cor, possibilitando assim o controle da resposta e a avaliação do êxito do aluno. (Candau, 1969, p. 39)

Com a ênfase na rápida resposta, Pressey teria antecipado a ideia da aprendizagem por reforço; no entanto, ainda como destaca Candau (1969), suas invenções não teriam acompanhado as teorias educacionais, as quais só viriam a se desenvolver posteriormente:

[D]esde um ponto de vista psicológico, as investigações no campo da aprendizagem ainda não se haviam desenvolvido adequadamente, ressentindo-se seus mecanismos [os mecanismos de Pressey] dessas deficiências: insistia-se mais na memorização que numa compreensão, a forma como se obtinha a resposta preocupava mais que a motivação. (Candau, 1969, p. 36; ressalva minha)

Vislumbrando uma “revolução industrial na educação” (Benjamin, 1988; Candau, 1969), Sidney Pressey acreditava nos desenvolvimentos tecnológicos como solução para incompetências típicas da “educação tradicional”, como cita Benjamin (1988)²⁷:

No capítulo final de seu livro de 1993 intitulado Psicologia e a Nova Educação, Pressey escreveu,
Deve haver uma “Revolução Industrial” na educação, em que a ciência educacional e a ingenuidade da tecnologia educacional sejam combinadas para modernizar os procedimentos grosseiramente ineficientes e desajeitados da educação convencional. O trabalho nas escolas do futuro será maravilhosamente organizado, ainda que de forma simples, para se ajustar quase que automaticamente às diferenças individuais e às características do processo de aprendizagem. Haverá vários esquemas e dispositivos de economia de trabalho, e até máquinas – nem todas para a

²⁷ Para um histórico detalhado de sucessos, fracassos, patentes etc na história das máquinas de ensinar, ver o artigo de Benjamin, “A History of Teaching Machines” (1988).

mecanização da educação, mas para libertar professor e aluno do trabalho enfadonho e da incompetência educacional (pp. 582- 583) (Benjamin, 1988, p. 707)²⁸.

A seguir, veremos que as máquinas de ensinar de Skinner surgiram em meio a uma atmosfera mais propensa, diga-se, do que as de Pressey, fator que pode ter contribuído para que ele conquistasse um maior reconhecimento em relação ao antecessor.

2.3.1

Ensino adaptativo: primórdios

Sobre uma das invenções de Pressey, B. F. Skinner, considerado o pai do behaviorismo, escreveu:

Ao confirmar as respostas corretas e ao enfraquecer as respostas que não deviam ter sido adquiridas, a máquina de autoavaliação, com efeito, ensina; mas não foi concebida primordialmente para este propósito. Não obstante, Pressey parece ter sido o primeiro a acentuar a importância do resultado (“feedback”) imediato na educação e a propor um sistema, no qual o aluno pode progredir no seu próprio ritmo. Viu a necessidade de equipamento básico para a realização destes objetivos. E, acima de tudo, concebeu uma máquina que (ao contrário dos recursos audiovisuais que começavam a ser desenvolvidos) permitia ao estudante uma participação ativa (Skinner, 1958, p. 30).

Segundo Skinner, para além de perceber a necessidade de individualizar o ensino, permitindo que cada aluno seguisse seu próprio ritmo, e de liberar o professor de tarefas repetitivas, Pressey teria levado em conta a importância de atribuir ao aluno um lugar ativo no processo de aprendizagem. A despeito disso, suas invenções teriam sido pouco reconhecidas; e a explicação poderia se dever ao momento em que elas emergiram – além de o mundo da educação não estar preparado para tais novidades, havia uma certa inércia cultural (Skinner, 1958). De fato, Pressey estava trabalhando para desenvolver suas máquinas de ensinar nos anos de 1930, enquanto Skinner lançou suas criações no final dos anos de 1950 e início da década de 1960. Sendo assim, se Pressey contou com a escassez do cenário

²⁸ No original: *In the final chapter of his 1933 book entitled Psychology and the New Education, Pressey wrote,*

There must be an "industrial revolution" in education, in which educational science and the ingenuity of educational technology combine to modernize the grossly inefficient and clumsy procedures of conventional education. Work in the schools of the future will be marvelously though simply organized, so as to adjust almost automatically to individual differences and the characteristics of the learning process. There will be many labor- savings schemes and devices, and even machines--not at all for the mechanizing of education, but freeing of teacher and pupil from educational drudgery and incompetence. (Benjamin, 1988, p. 707)

da crise de 1929²⁹, em que não havia demanda por um aumento do ritmo da educação, Skinner teve como benefício o embalo da automação da educação pública que emergiu com a Segunda Guerra Mundial. Na guerra, como relata Benjamin (1988), filmes foram usados para treinar milhares de militares, o que alavancou o sucesso dos recursos audiovisuais nos sistemas públicos de ensino. A primeira estação de televisão norte-americana foi inaugurada em 1953 e, dez anos depois, já havia 60 afiliadas à National Educational Television Network. Para completar o cenário, os Estados Unidos sofriam, no final dos anos de 1950, com falta de professores, “o que prometia ser exacerbado com o baby-boom pós Segunda Guerra” (Lucy, 1960, apud Benjamin, 1988). Não fosse suficiente, os soviéticos lançaram seu Sputnik – o que teria contribuído para pôr em xeque a qualidade da educação dos EUA, especialmente nos campos da ciência e da matemática. Esses acontecimentos é que aparentemente foram motivos para que as invenções de Skinner encontrassem um terreno mais receptivo do que seu antecessor. Com as promessas – que seguiam firmes – de ensino mais rápido e de alívio de certos fardos das costas dos professores, havia mais de sessenta máquinas de ensinar disponíveis no mercado norte-americano em 1962, segundo Benjamin (1988).

Foi a partir de algo que observou na escola de sua filha que Skinner desenvolveu uma máquina para ajudar a ensinar aritmética. Ele notou que todos os estudantes precisavam seguir o mesmo ritmo e que tinham de esperar 24 horas para obter retorno sobre seu desempenho, o que o incomodou. Então, desenvolveu um

²⁹ Segundo Watters (2021), a Grande Depressão arruinou os planos de Pressey, mas, mesmo que o mercado não quebrasse, ele teria tido dificuldade em persuadir a indústria das máquinas de escrever a construírem sua máquina de ensinar. Afinal, os fabricantes de máquinas de escrever estavam convencidos de que elas seriam o dispositivo educacional moderno ideal. Ela cita Ben D. Wood, professor de psicologia educacional da Columbia University, como uma figura que teve um papel definitivo para que a indústria apostasse nisso. Em 1928, os fabricantes das empresas de máquinas de escrever Remington Typewriter, Royal Typewriter, Smith & Corona e Underwood Typewriter fizeram uma reunião em que formaram um “bureau de educação” para “obter fatos e evidência experimental das vantagens do uso de máquina de escrever portátil como um instrumento educacional” (Watters, 2021, p. 62). Em 1929 foi lançado o Educational Bureau of Portable Typewriters, tendo sido Ben Wood convidado a liderar a pesquisa sobre a efetividade do uso desses dispositivos na sala de aula. Apesar de o conglomerado formado pelos fabricantes ter se comprometido em não fazer publicidade ou marketing de seus produtos diretamente a escolas que participassem de seus experimentos, Watters relata que houve inúmeros episódios de contatos comerciais com as escolas, já que não eram implementadas sanções significativas às companhias por essa sua atitude. “No doubt, what the typewriter manufacturers were interested in was less ‘the integrity of the experiment’ and more the benefits – the *commercial* benefits, to be clear – they might accrue from what was a significant financial investment in the bureau” (Watters, 2021, p. 63).

dispositivo que apresentava um problema, uma resposta e uma luz que piscava quando a resposta fornecida estava correta (Benjamin, 1988, p. 708). Estaria resolvendo, assim, o problema de ritmo individual de cada aluno. Vale ressaltar, porém, que, se o que incomodou Skinner e fez emergir suas invenções foi parecido com o que Pressey sentiu, as ideias envolvidas nas criações de um e de outro tinham diferenças relevantes.

Enquanto Pressey desenvolveu dispositivos pensando em testar os alunos – de modo que, antes de utilizar as máquinas, eles deveriam já ter estudado as lições – o behaviorista Skinner voltou-se para as possibilidades de realmente ensinar conteúdos novos aos alunos por intermédio de tais tecnologias. Como coloca Benjamin (1988), Skinner acreditava que a aprendizagem ocorreria em pequenos passos e que o material a ser disponibilizado deveria ser construído de uma maneira coerente com um repertório responsivo, passo a passo. O termo “ensino programado”, objeto de estudo de Candau (1969), teria sido cunhado por Skinner para descrever a “informação construída dessa maneira lógica e sistemática” (Benjamin, 1988, p. 708). Nas máquinas de Skinner – as quais não se limitaram à citada invenção para ensinar matemática, mas adquiriram contornos diversos – o aluno foi convocado a escrever suas próprias respostas, podendo então compará-las com as que os dispositivos ofereciam; o processo representava uma saída do modelo de múltipla escolha para um discursivo.

Como descrevem Holmes, Bialik e Fadel (2019, p. 95-96), a máquina que Skinner lançou em 1958 com base em seus experimentos para condicionar ratos e pombos era uma caixa de madeira com furos ou janelas na tampa. Perguntas escritas em discos de papel apareciam em um desses furos, ou janelas, enquanto o estudante deveria escrever uma resposta num rolo de papel que ficava visível por meio de outra janela, também usada para o professor fazer marcações depois. O mecanismo automaticamente cobria a resposta do aluno, que não poderia ser modificada depois de feita, e revelava o conteúdo da resposta correta. Ainda segundo os autores, porém, por mais que tenha se voltado para máquinas que pudessem gerar aos estudantes novos conteúdos e que os permitissem escrever suas próprias respostas em vez de selecioná-las em um universo pré-determinado, as criações de Skinner não podem ser consideradas *adaptativas*. Isto porque não forneciam nenhum tipo de acomodação das questões ou de sua ordem de exibição em função do desempenho do estudante que estivesse usando a máquina. Por mais que os

estudantes pudessem prosseguir segundo seu próprio ritmo, todos passavam pela mesma lista de questões dos demais e tinham acesso às questões na mesma ordem. Uma máquina adaptativa, em teoria, *se adapta* ao ritmo do estudante que a está utilizando e customiza a entrega de conteúdo para ele.

Holmes, Bialik e Fadel (2019, p. 97) apontam Norman Crowder como um precursor do ensino adaptativo, enquanto Gordon Pask teria desenvolvido provavelmente a primeira máquina de ensinar verdadeiramente adaptativa no início dos anos de 1950. Crowder desenvolveu uma alternativa às máquinas de ensinar lineares, como as de Skinner, que os autores classificam como instrução programada intrínseca ou ramificada – em inglês, *intrinsic* ou *branching programming* (ver também Candau, 1969, p. 61). Trabalhando como psicólogo na Força Aérea dos EUA, ele estava interessado em treinar engenheiros para que se tornassem capazes de detectar problemas de funcionamento em equipamentos eletrônicos.

O sistema de Crowder, como relatam Holmes, Bialik e Fadel (2019), foi projetado de maneira tal que o sujeito a ser treinado era apresentado a uma página com informações seguida de questões de múltipla escolha; cada resposta possível conduzia o estudante a uma nova página. Se o aluno, então, escolhesse a resposta errada, encontraria posteriormente conteúdo voltado para corrigir o que ele não havia entendido bem; já se acertasse a resposta, encontraria em seguida uma página com mais conteúdos relacionados e que pudessem expandir seus conhecimentos. A ideia, enfim, era que os estudantes seguissem jornadas de aprendizagem diferentes dependendo do que fossem acertando ou errando, algo semelhante ao que se observa no conceito de trilhas de aprendizagem, hoje aplicado a soluções online de ensino. Desse modo, pode-se imaginar uma linha reta a ser perseguida pelos estudantes numa máquina de ensinar como as que Skinner criou, ao passo que, para uma máquina baseada em programação intrínseca, a melhor metáfora talvez fosse a de uma árvore com seus múltiplos galhos. Candau (1969) destaca o aspecto adaptativo da invenção de Crowder, contrastando-o com a linearidade típica da máquina de ensinar de Skinner, da seguinte maneira:

A programação intrínseca diferencia-se basicamente da programação de tipo Skinneriano, pois, enquanto esta é linear e apresenta o material na mesma sequência para todos os alunos, a programação intrínseca procura adaptar mais o conteúdo e a sequência do material apresentado às dificuldades próprias de cada aluno. (Candau, 1969, p. 64).

Citando Crowder, ela faz referência ainda ao fato de que as máquinas baseadas em programação intrínseca prescindiam de um tutor externo para fornecer as próximas lições aos alunos – necessidade indispensável no caso de máquinas lineares:

“A expressão ‘programação intrínseca’ refere-se simplesmente ao fato de que o programa necessário de alternativas constitui-se no interior do próprio material de tal modo que nenhum mecanismo exterior ao programa é necessário”. (Candau, 1969, p. 64)

Uma ressalva vale ser feita, a partir de observação de Candau sobre as premissas por trás da máquina adaptativa de Crowder: ao contrário do que acontecia com Skinner, o interesse dele, ao criar tal dispositivo, voltava-se mais para a eficácia da comunicação dirigida aos estudantes quanto ao treinamento realizado do que à preocupação em seguir alguma teoria de aprendizagem. Trata-se de algo a ser levado em consideração, especialmente quando se observa que Skinner, para além de desenvolver dispositivos para ensinar, conduzia pesquisas dedicadas à compreensão da aprendizagem. Já Pask, um cibernético britânico, desenvolveu e patenteou em 1956 um recurso conhecido como SAKI (*Self-Organizing Automatic Keyboard Instructor*): um teclado autoadaptativo projetado para treinar operadores de cartões perfurados, que pode – esse sim, segundo Holmes, Bialik e Fadel (2019), Watters (2015)³⁰ e Candau (1969) – ser considerado um sistema adaptativo pioneiro³¹. Basicamente porque as tarefas eram diferentes para cada aluno e, de acordo com seu desempenho, operavam a partir de um modelo probabilístico. Segundo Watters (2015),

Como muitas máquinas de ensinar (naquela época e agora), a SAKI pretendia funcionar como um tutor humano. Porém, ao contrário das máquinas de ensino anteriores, o componente adaptativo dos dispositivos de Pask oferece mais do que apenas uma avaliação do certo ou errado: eles identificam e medem as respostas de um aluno – precisão, tempo de resposta – e ajustam a próxima pergunta de acordo. Ou seja, a dificuldade das questões não é pré-programada ou pré-ordenada (...) A máquina responde *algoritmicamente*. (Watters, 2015, não paginado)³².

³⁰ <http://hackededucation.com/2015/03/28/pask> (Acesso em 10 de dezembro de 2020).

³¹ Há um esquema mostrando como o sistema funciona em <http://hackededucation.com/2015/03/28/pask>

³² No original: *Like many teaching machines (then and now), SAKI purported to function like a human tutor. But unlike earlier teaching machines, the adaptive component of Pask's devices offers more than just an assessment of right or wrong: they identify and measure a student's answers – accuracy, response time – and adjust the next question accordingly. That is, the difficulty of the questions are not pre-programmed or pre-ordained. (...) The machine responds **algorithmically**. (Watters, 2015).*

Sobre Pask e a SAKI, Candau (1969) escreveu:

Segundo ele, era impossível projetar um programa adequado para ensinar estas habilidades [habilidades manuais] sem levar em consideração as mudanças de atitude do aluno e os períodos de interesse e fadiga que se sucedem durante o processo de aprendizagem. Como resultado dessas conclusões, interessou-se por máquinas que “aprendem” o padrão de conduta dos alunos e constroem os seus próprios programas com base nesta informação. Ao regressar a seu país inventou, juntamente com Christopher Bailey e McKinnon Wood, uma máquina de ensinar adaptativa chamada SAKI (...) especialmente desenhada para exercitar futuras perfuradoras de fichas. (Candau, 1969, p. 74).

À semelhança do que ocorre na história da IA em geral, que conta com longos períodos de inverno entremeados por momentos de excitação, a história das máquinas de ensinar é permeada por pausas e hesitações. Na seção a seguir abordo as desconfiças que surgiram a partir das primeiras máquinas de ensinar, por fim colocando-as em xeque.

2.3.2

Derrocada das (primeiras) máquinas de ensinar

Na década de 1960, era notável o aumento do interesse pelas máquinas de ensinar, capitaneadas por Skinner e suas teorias de aprendizagem. Tanto o campo educacional quanto a indústria e os militares se interessavam pelas aplicações e a sua utilidade em treinamentos (Benjamin, 1988). Porém, já naquela época ecoavam enxurradas de críticas e questionamentos acerca dos usos de tais recursos e as possibilidades de um futuro distópico liderado por eles, *a la* George Orwell e Aldous Huxley³³. Parte dessas preocupações podiam ser identificadas pelos títulos de artigos lançados então, como:

“As pessoas podem ser ensinadas como pombos?” (Boehm, 1960); “Podem as máquinas substituir professores?” (Luce, 1960); “Máquinas de Ensinar – Bênção ou Maldição?” (K. Gilmore, 1961); “Robôs vão ensinar seus filhos?” (Bell, 1961); “Máquinas de Ensinar realmente ensinam?” (Margolis, 1963); “Qual é este? O Novo Mundo das Máquinas de Ensinar ou o Admirável Mundo Novo das Máquinas de Ensinar?” (Morello, 1965). (Benjamin, 1988)³⁴.

³³ A ficção sobre a vigilância excessiva da sociedade, intitulada “1984”, de George Orwell, foi lançada em 1945; já o romance “Admirável Mundo Novo”, de Aldous Huxley, havia sido lançado anos antes, em 1932.

³⁴ No original: “Can People Be Taught Like Pigeons?” (Boehm, 1960); “Can Machines Replace Teachers?” (Luce, 1960); “Teaching Machines --- Blessing or Curse?” (K. Gilmore, 1961); “Will Robots Teach Your Children?” (Bell, 1961); “Do Teaching Machines Really Teach?” (Margolis,

Associada à ideia de desumanização da educação que marca essas manchetes estava a suposição de que, em vez de contribuir para individualizar o ensino, colaborando para o respeito ao ritmo de aprendizagem de cada aluno, as máquinas de ensinar acabariam por levar os professores a dar menos atenção aos alunos em sala de aula – um efeito completamente oposto às alardeadas propostas iniciais de personalização e individualização. Outro alvo do olhar dos críticos, entre eles pais e educadores, era a dúvida quanto a tais tecnologias poderem ou não ensinar, de fato (Benjamin, 1989):

Skinner pensava que qualquer coisa que pudesse ser verbalizada poderia ser ensinada por uma máquina de ensinar ("The Critics Speak," 1967), mas outros sentiam que os programas somente operariam bem para aquelas disciplinas facilmente fragmentadas, como por exemplo as línguas estrangeiras e a matemática. Um crítico notou que talvez fosse possível aprender um poema usando uma dessas máquinas, mas a máquina não poderia ensinar o amor pela poesia (Margolis, 1963, apud Benjamin, 1989)^{35, 36}.

Além das questões de se poder ou não realmente ensinar usando tais tecnologias, do conteúdo ou dos limites daquilo que se poderia ensinar, Benjamin (1989, p. 709-710) ressalta, ainda, o temor de alguns quanto ao fato de os professores não saberem usar tais tecnologias, o que poderia fazer com que seu potencial fosse subaproveitado; em outro extremo, havia o receio de que os dispositivos ensinassem até bem demais, o que deixaria os aprendizes nas mãos daqueles que fossem capazes de produzir as máquinas a serem aplicadas para o ensino programado. Ela menciona o medo da perda do emprego, por parte dos professores, e afirma que alguns administradores escolares viram no uso das

1963); *"Which Is It? New World of Teaching Machines or Brave New Teaching Machines?"* (Morello, 1965). (Benjamin, 1988).

³⁵ Skinner argued that anything that could be verbalized could be taught by a teaching machine ("The Critics Speak," 1967), but others felt the programs worked well only for those subjects that were easily fragmented, for example, foreign languages and mathematics. One critic noted that it might be possible to learn a poem using one of the machines, but the machine could not teach the love for poetry (Margolis, 1963).

³⁶ Este trecho é particularmente interessante e pertinente para esta pesquisa porque, com frequência, uma das críticas dirigida às abordagens da cognição enativa e corporificada, entre outras dos 4Es, é a impossibilidade de que essas abordagens expliquem funções cognitivas superiores, estando limitadas a explicar as inferiores ou mais básicas. Mas o que essas abordagens buscam não é somente compreender as funções cognitivas básicas, mas o raciocínio também – porém revestindo-o de uma camada que inclui o corpo, a estética, a experiência, as emoções – aspectos que vou trabalhar no decorrer da tese. O ponto é que talvez o que se compreende como “fragmentado”, ou passível de fragmentação, quando se tem uma concepção da razão que é apartada das emoções, por exemplo; ou quando se parte de uma concepção de cognição que divide a operação cognitiva em partes para depois reuni-las num todo e vez de prezar pela apreensão de uma totalidade, pode-se conceber a aprendizagem também de maneira fragmentada, separando disciplinas que teriam maiores ou menores possibilidades de fragmentação.

máquinas uma desculpa para aumentar a quantidade de alunos atendidos por professor em sala de aula. Um crítico teria apelidado as máquinas de meras “viradoras de páginas caras”, querendo argumentar que tais dispositivos seriam recursos desnecessários que não fariam nada mais do que aquilo que já se podia fazer sem eles; a isso, Skinner e outros teriam respondido defendendo que os recursos poderiam evitar que alunos *colassem* as respostas, além de proporcionar uma melhor apresentação dos conteúdos.

Em 1961, um estatuto intitulado “Self-Instructional Materials and Devices” foi lançado pela American Educational Research Association, a National Education Association e a APA – American Psychological Association (ainda que psicólogos estivessem divididos quanto às tecnologias) com a intenção de chamar a atenção para os cuidados necessários quanto às máquinas de ensinar. Por mais que Benjamin (1989) apenas mencione brevemente a questão de os sistemas guardarem informações dos alunos, um dos pontos do estatuto (o sexto) voltava-se justamente para a questão dos dados armazenados, algo que hoje está muito em voga na forma de discussões sobre *big data* na educação:

Uma característica importante de quase todos os materiais autoinstrucionais é o registro das respostas do estudante, que fornece uma base para a revisão do programa. O potencial comprador deve questionar até que ponto a revisão foi baseada na resposta do estudante e como os testes preliminares foram conduzidos. (APA, 1961, p. 512)³⁷

Benjamin (1989) relata que, ao final dos anos de 1960, a maior parte das máquinas de ensinar havia sido retirada do mercado, enquanto os artigos sobre o tema começaram a sumir dos periódicos. A atenção daqueles que se interessavam pelas tecnologias educacionais passou a se voltar para o ensino apoiado por computadores (CAI – *Computer-Assisted Instruction*). Escrevendo em 1989, Benjamin destaca que a chegada dos computadores às salas de aula foi acompanhada por receios similares aos que acompanharam a queda da popularização das máquinas de ensinar de Pressey, Skinner e outros.

Talvez a inércia cultural a que Skinner se referia na época de Pressey ainda exista hoje, pelo menos no que diz respeito ao sistema educacional [norte]americano. Nos

³⁷ No original: *An important feature of almost all self-instructional materials is the record of the student's responses which provides a basis for revising the program. The prospective purchaser should ask about the extent to which revision has been based on student response and how the preliminary tryout was conducted.* (APA, 1961, p. 512)

últimos 50 anos, uma série de novas abordagens prometeram suplantam o professor ou, pelo menos, alterar substancialmente o papel do professor na sala de aula. Primeiro foram os filmes, depois as máquinas de ensino e instrução programada, depois a televisão, depois os cursos de sistema de instrução personalizado (PSI) e agora os computadores. No entanto, pode-se argumentar que nenhuma dessas tecnologias sobreviveu ao seu faturamento avançado. Todos eles desempenharam um papel na educação e continuam a fazê-lo, mas é um papel secundário (...) se o comportamento passado prevê o comportamento futuro, então parece improvável que os computadores ou qualquer outra máquina de ensino desempenhem mais do que um papel de apoio em sala de aula. (Benjamin, 1989, p. 711)³⁸

Será mesmo que é possível estabelecer previsões adequadas com base no que aconteceu no passado, como Benjamin coloca na última frase da citação acima? A IA baseada na aprendizagem de máquina pode apostar que sim, já que é fundamentada em previsões que – ao menos assim se supõe – se tornam mais precisas na medida em que a máquina “aprende”. Há, porém, uma espécie de marco na linha do tempo das máquinas de ensinar na educação que se caracteriza justamente pelo momento em que elas se tornam *máquinas de aprender*. Parceiros do *big data*, os sistemas de machine learning recolocam a automação da aprendizagem no cenário educacional, mas talvez propondo novos questionamentos. A preocupação da APA na década de 1960, no entanto, ainda é (com a devida atualização) uma preocupação hoje: a gravação e o registro dos dados dos estudantes levam a problemas de privacidade e questões éticas sensíveis. A imensa capacidade de processamento de dados dos sistemas de machine learning acrescenta uma camada considerável a essa preocupação, que envolve os usos que se faz desses dados, os impactos para a vida dos que têm seus dados expostos, e qual seria a extensão possível desses impactos.

³⁸ No original: *Perhaps the cultural inertia that Skinner referred to in Pressey's time still exists today, at least with regard to the American educational system. Over the past 50 years a number of new approaches have promised to supplant the teacher or at least substantially alter the teacher's role in the classroom. First it was movies, then teaching machines and programmed instruction, then television, then personalized system of instruction (PSI) courses, and now computers. Yet it can be argued that none of these technologies has lived up to its advanced billing. They have all played a role in education and continue to do so, but it is a minor role (...) if past behavior is a predictor of future behavior, then it seems unlikely that computers or any other teaching machines will play more than a supporting role in the classroom. (Benjamin, 1989, p. 711)*

2.4

Máquinas de ensinar ontem e hoje

As antigas máquinas de ensinar eram, sem dúvida, diferentes das plataformas atuais baseadas em machine learning. No entanto, os fatores que impulsionam os usos de ML na educação, na atualidade, são muito semelhantes a aqueles que foram associados aos usos das primeiras máquinas de ensinar, analógicas, décadas atrás. A teoria de aprendizagem predominante na época das máquinas analógicas era o behaviorismo, que vem sendo associado por estudiosos do tema às máquinas atuais. É preciso destacar que não apenas as vantagens supostamente associadas, mas os pressupostos por trás das primeiras máquinas de ensinar mostram-se bastante parecidos com os que estão por trás dos sistemas das plataformas educacionais baseadas em aprendizagem de máquina. E esses pressupostos podem revelar algo importante sobre aquilo que se entende por ensino e aprendizagem, quando se admite o uso de tais recursos – o que se torna especialmente relevante para uma investigação sobre fundamentos e premissas, como uma tese em filosofia da educação motiva-se a fazer.

Para além dos benefícios associados à IA na forma de machine learning na educação, e que se assemelham às vantagens defendidas para as primeiras máquinas de ensinar – como a personalização, a suposta facilidade de compreender e atender melhor às dificuldades dos alunos e uma liberdade maior para o professor – há nas entrelinhas um discurso bastante dúbio, e que suscita críticas parecidas com aquelas que foram direcionadas às primeiras máquinas de ensinar quando elas entraram em seu “inverno” após uma fase de euforia. Se, por um lado, parece haver preocupação com o professor, sua carga de trabalho e as dificuldades inerentes ao fato de que lida com alunos bastante diversos (sem mencionar as quantidades cada vez maiores de alunos por turma), por vezes – por exemplo – um *chatbot* é colocado como mais disponível do que o professor para responder a perguntas do alunado, gerando uma experiência supostamente melhor para os estudantes em sua interação com tais sistemas. Como se vê, nesses casos aparentemente a hipótese de que se parte é a de que as máquinas são incansáveis, mais interessantes e sedutoras e, em última instância, mais eficientes. Por que é possível que esse discurso esteja crescendo?

Segundo Watters (2021), um problema é que a indústria de tecnologia sempre financiou projetos de pesquisa em educação com a intenção não de que fossem

feitas descobertas espontâneas – como deve ser uma pesquisa de fato – mas de que servissem para corroborar a lógica de suas invenções e, em última instância, vender os artefatos desenvolvidos. Não à toa, como Watters relata, o então professor de Psicologia Educacional da Columbia University Ben Wood foi contratado por fabricantes de máquinas de escrever no final dos anos de 1920. Wood é autor de um estudo, muito citado, em que defendia que o uso de máquinas de escrever em sala de aula seria capaz de desenvolver os hábitos de leitura dos estudantes e aumentar quantidade de produção textual que eles poderiam fazer. Ainda segundo a autora, em 1928, George Eastman, cofundador da Kodak, teria financiado Wood e Frank Freeman (um empresário do ramo das tecnologias educacionais) para que "investigassem" e, idealmente, revelassem as vantagens educacionais do uso do filme. Conduzindo um estudo com cerca de onze mil estudantes, procuraram então comprovar que era efetivo usar vídeos para ensinar geografia e ciências, a partir da aplicação imediata de testes ao final da apresentação do material visual, mostrando ainda que esses resultados eram melhores do que quando se aplicava o uso de materiais impressos. Além da (eticamente conflitante) associação entre indústria e pesquisas (conduzidas desde a partida sob vieses e para comprovar teses, em vez de serem de fato investigações), um processo de retroalimentação que acaba por priorizar o valor comercial sobre o valor educacional, vale destacar que as tecnologias educacionais desenvolvidas sob a égide da personalização ou individualização do ensino têm um traço marcante: são tecnologias que, basicamente, aplicam testes. A teoria de aprendizagem predominante na época das máquinas analógicas era o behaviorismo, que vem sendo associado por estudiosos do tema às máquinas atuais – voltarei a isso adiante neste capítulo.

Como Watters (2021) conta, o conhecido exame New York Regents Examination, utilizado até hoje para avaliar alunos do ensino médio nas disciplinas básicas, era administrado por uma empresa pertencente a Wood, que depois assumiu um projeto ainda maior nos EUA, o Pennsylvania Study, envolvendo quase trinta mil estudantes do ensino médio e quase cinco mil de graduação. Wood buscou parceiros que pudessem automatizar as correções dos testes; as empresas não lhe deram muita atenção, exceto a IBM, na figura de seu diretor, Thomas J. Watson. Junto à IBM, Wood criou dispositivos capazes de tabular os testes. Wood (Watters, 2021, p. 69) defendia que a individualização da educação exigiria que professores conhecessem os estudantes, e que o grande desafio não era ensinar, mas aprender

sobre os alunos. No entanto, conhecer os alunos, para Wood, significava testá-los, fosse via conteúdo, fosse via testes de personalidade, de inteligência e de aptidão ou por meio de análises psicológicas. Indo mais além, Wood acreditava que as

testagens por si só não dariam conta disso, e que as práticas pedagógicas precisariam mudar para responder a essa necessidade de dados. Em 1934, ele escreveu que seria necessário construir uma filosofia e um sistema que fossem

adequados para usar os resultados dos testes para construir uma orientação educacional, e que esse sistema demandaria um maquinário específico. Não por menos, Watters (2021) defende que a individualização e a personalização na educação, no final das contas, há muito acabam se resumindo a uma coleta de dados.

Sendo assim, argumento que, ao menos em parte, isso pode se dar porque a ideia de aprendizagem que serve como base para esse processo de automatização é limitada e/ou limitante. Afinal, implícitas no discurso de ensinar melhor com máquinas existem ideias como a valorização da capacidade de processamento de informações (seria por isso que se defenderia a eficiência de um *chatbot* em detrimento da capacidade de um professor em certos casos?); a rapidez como sinônimo de efetividade; um afastamento das emoções dos processos de aprendizagem; a valorização do cumprimento de mais tarefas em menos tempo e até mesmo um julgamento daquilo que seria mais ou menos digno de receber a atenção e contar com a dedicação e o esforço de professores e alunos. Um dos pontos incoerentes é o seguinte: se a IA é mesmo tão valorizada e as técnicas de machine learning são tão eficientes, elas poderiam lidar muito bem com tarefas como a de responder a perguntas; mas, ao mesmo tempo, perguntas de alunos não parecem tão importantes, na verdade acabam por perturbar o professor e ocupá-lo sem necessidade. Isso contradiz a própria função que essas tecnologias teriam.

Como se vê, o discurso predominante assume uma série de premissas como certas e esconde vários possíveis desalinhamentos. Assim, as perguntas que surgem – como mencionei – acabam sendo muito parecidas com aquelas que os críticos da educação fizeram diante das antigas máquinas usadas no ensino. As questões levantadas nos títulos de artigos mencionados por Benjamin (1988) e que reproduzi neste capítulo parecem ser ainda muito pertinentes, apesar de haver ingredientes novos nessa mistura, sendo um deles o *big data*. Críticos das tecnologias educacionais apontam questões ligadas, por exemplo, à privacidade dos alunos e professores, o que suscita pertinentes problemas éticos; questões relativas à

aprendizagem, como a pertinência daquilo que se deseja ensinar por meio de plataformas de ensino – se elas não estariam entrando para preencher lacunas alegadas justamente para justificá-las, mas que talvez nem existam (Bannell, 2017) – e, ainda, se de fato essas plataformas teriam maior capacidade de atender aos alunos individualmente ou se, no fim das contas, não estariam contribuindo para pasteurizar ainda mais o processo de ensino, tornando possível o atendimento de mais alunos em menos tempo em nome da eficiência (*eficiência* nesse sentido também é discutida em Knox, Williamson e Bayne, 2019), terminando por sobrecarregar o professor, em vez de livrá-lo de tarefas ditas maçantes. Mais uma vez, nota-se preocupações semelhantes às que foram direcionadas pelos críticos às primeiras máquinas de ensinar.

Um olhar atento para a aprendizagem de máquina no contexto educacional pode ainda eliminar uma suposta aura neutra envolvendo os dados para concebê-los como uma consequência da própria maneira como os sistemas foram construídos. Isso mitiga a ingenuidade de conceber dados como algo espontâneo que simplesmente aparece e pode então ser capturado para determinados fins. Explicando de outra maneira: a hipótese é a de que sistemas de ML seriam desenvolvidos de uma forma tal que leva os alunos a demonstrar certos resultados, dentro de uma gama de resultados possíveis, em detrimento de outros; isto é, os dados não seriam gerados de maneira assim tão espontânea. Não haveria surpresa nesse processo, ou haveria, digamos, uma limitada abertura para o inesperado, dentro de um cenário já previsto; as máquinas “aprenderiam” dentro de um percurso já pré-estabelecido, e os alunos apenas seguiriam uma linha anteriormente traçada³⁹.

Em uma crítica que estende esse aspecto ao extremo, Knox, Williamson e Bayne (2019) apontam uma forte tendência behaviorista em machine learning aplicado à educação, um conceito que chamam de “machine behaviourism”: combinações de teorias radicalmente behavioristas com sistemas de machine learning estariam trabalhando contra noções de autonomia e participação dos estudantes, procurando intervir na conduta educacional e moldando o comportamento dos alunos em direção a objetivos predefinidos. Ainda segundo os

³⁹ Um percurso pré-estabelecido para a aprendizagem seria algo radicalmente contra a proposta enativista para a cognição, já que segundo os enativistas o caminho é construído justamente durante o caminhar. Isso será desenvolvido no decorrer da tese, especialmente no capítulo 4.

autores (2019, p. 1-2), métodos associados à ciência de dados estão envolvidos na reformulação, conceitual e material, daquilo que “se tornou a principal preocupação da educação contemporânea: a aprendizagem”.

Essa preocupação central com a aprendizagem, no contexto de que falam Knox, Williamson e Bayne (2019), tem uma conotação específica. Aproxima-se do conceito de “learnification” (Biesta, 2005), o qual encapsula a aprendizagem como algo que pode ser consumido, e o estudante como o consumidor. Dentro dessa lógica, o estudante precisaria ser atendido em suas necessidades de aprendizagem, as quais já lhe seriam claras – o que, como os autores apontam, é problemático na raiz, pois seria justamente o processo de aprendizagem que possibilitaria que os estudantes compreendessem as suas necessidades de aprendizagem.

A primeira razão, portanto, para ser contra a aprendizagem – ou seja, ser contra uma expressão que torna possível apresentar a educação em termos de “atender às necessidades do aprendiz” – é que a suposição subjacente de que os aprendizes vêm à educação já com uma compreensão clara de quais são suas necessidades, é uma suposição altamente questionável. Ela interpreta mal o papel e a posição do profissional educacional no processo, e o papel e a posição do aprendiz. Esquece que uma das principais razões para se envolver na educação é precisamente descobrir do que é que realmente o indivíduo precisa – um processo no qual os profissionais da educação desempenham um papel crucial porque a maior parte de sua experiência está precisamente ali. (Biesta, 2005, p. 59)⁴⁰

Isto é, antes de passar pelo processo, pela experiência de aprender, aprendizes não teriam como conhecer suas necessidades e demandas – muito menos ser capaz de exigi-las de instituições de ensino como se exige um outro produto qualquer, pré-fabricado. Seria de responsabilidade das instituições definir objetivos educacionais e prezar pelo seu papel nesses processos, de modo a evitar que a função dos professores pudesse ser considerada meramente responsiva e que a instituição passasse a existir estritamente para dar conta de necessidades dos alunos-consumidores. *Learnification* seria, assim, um fenômeno cego a questões mais amplas sobre o papel e o propósito da educação na sociedade de um modo geral, sobre como metas e aspirações são negociados e estabelecidos, e quais os

⁴⁰ No original: *The first reason, therefore, to be against learning – that is, to be against a language which makes it possible to present education in terms of «meeting the needs of the learner» – is that the underlying assumption that learners come to education with a clear understanding of what their needs are, is a highly questionable assumption. It both misconstrues the role and position of the educational professional in the process, and the role and position of the learner. It forgets that a major reason for engaging in education is precisely to find out what it is that one actually needs – a process in which educational professionals play a crucial role because a major part of their expertise lies precisely there. (Biesta, 2005, p. 59)*

tipos de estruturas de poder que sustentam esses processos; a ideia é de que, quando se reduz a educação à aprendizagem, nesse sentido do conceito, a educação fica restrita a uma relação de consumo⁴¹.

O processo que os autores apontam como “datafication”, por sua vez, mais uma vez alteraria o controle do processo de aprendizagem, deslocando-o do aluno para os detentores dos dados (como também ressaltado por Benjamin, 1989, acerca das máquinas de ensinar antigas). A entrada massiva da ciência de dados na educação não se dá ao acaso, mas como resultado de um cenário propício, em que o aprendiz teoricamente desempenha um papel central. Os dados desse aluno seriam recolhidos e analisados por sistemas capazes de dar conta desses dados de uma maneira que, alegadamente, somente sistemas com tal potência técnica poderiam, assim adentrando um universo e desempenhando uma tarefa na qual humanos não poderiam ser bem-sucedidos. Ainda que seja verdade que sistemas de machine learning tenham uma capacidade de processamento de dados muito superior à humana, há aí um nó (ou vários). Numa forte crítica à tecnologia como suposta detentora das soluções para as lacunas da educação, Knox, Williamson e Bayne (2019) ressaltam que, por trás da ideia de que haja um universo desconhecido para educadores (o que levaria à marginalização da experiência do educador), existe a pressuposição de uma falha na compreensão que educadores teriam e que poderia somente ser preenchida ou esclarecida com dados (*big data*). Nesse contexto, os dados teriam grande poder; conseqüentemente,

[N]ão somente os dados seriam posicionados antes dos desejos do aprendiz, como a fonte autoritária para a ação educacional, mas o papel do estudante em si seria reposicionado como produto das tecnologias analíticas consumistas (Knox, Williamson & Bayne, 2019, p. 5).⁴²

⁴¹ Outro ponto é a questão da queda do Estado de bem-estar social, que provocou uma mudança na relação entre o Estado e o cidadão: ela se tornou uma relação em que o Estado é provedor de serviços públicos e quem paga imposto consome esses serviços. Na educação, o impacto disso é a accountability, o processo pelo qual aumenta a inspeção desses serviços e os protocolos de educação se tornam mais prescritivos. Essa lógica também resulta em pais que se veem como consumidores da educação de seus filhos e que entendem então poder decidir o que deve acontecer nas escolas. “This way of thinking introduces a logic which focuses on the users or consumers of the educational provision and a very suitable name for the consumer of education is, of course, «the learner»”. (Biesta, 2005, p. 57)

⁴² No original: [n]ot only is data positioned before the desires of the learner as the authoritative source for educational action, but the role of the learner itself is also recast as the product of consumerist analytic technologies (Knox, Williamson & Bayne, 2019, p. 5).

Isso se torna mais grave quando se concebe que os dados, em vez de serem produtos espontâneos de uma relação de uso humano da máquina, são gerados a partir de uma engenharia desenhada justamente para que emergjam já com certas características – o que os autores chamam, como já mencionado, de *machine behaviourism*; no contexto de que eles falam, “behaviour” deve ser entendido de maneira radical, como uma intervenção proposital; ou seja, a tarefa não é apenas observar o *comportamento* do estudante, posteriormente analisando-o, mas ativamente modificar e determinar a direção em que vai esse comportamento. Tal artimanha se implementaria, especificamente, em sistemas baseados em *reinforcement learning*. O aprendiz é tomado como um sujeito irracional, cujos comportamentos e ações podem não somente ser lidos por algoritmos de machine learning como modificados, reforçados ou desestimulados por essas plataformas; é desse modo que, como defendem os autores, o “machine behaviourism” marcaria um retorno à influência da psicologia behaviorista na prática educacional. Haveria uma forma de agir tomada como correta, e que já seria estabelecida de antemão. Os estudantes seriam conduzidos, empurrados (há uma expressão em inglês para isso, “nudge”)⁴³ na direção desses comportamentos esperados (Knox, Williamson & Bayne, 2019, p. 5)⁴⁴. Nada disso, vale comentar, parece ser muito diferente do que Watters apontou como resultante da contratação de investigadores para fazer pesquisas que apenas pudessem reforçar o que a indústria da tecnologia desejava divulgar e defender. Se um sistema já for concebido para seguir determinada direção, e assim influenciar e possivelmente limitar a atuação de quem o utiliza, não seria esse fator conflitante com a espontaneidade típica do processo de aprender? Esta é uma questão que precisa ser analisada quando o que se deseja é compreender de que maneiras as tecnologias digitais podem fazer parte dos

⁴³ Independentemente de haver consenso ou não a respeito da prática de *nudging* e das implicações éticas envolvida nisso, na minha perspectiva a aprendizagem de máquina parece restringir as opções de quem utiliza esses sistemas, quando se considera a cognição de uma maneira mais ampla do que as máquinas consideram; desenvolverei melhor esta ideia em capítulo mais à frente.

⁴⁴ Um exemplo de *nudging*, dado pelos autores, e que torna mais fácil a compreensão prática de uma implementação dessa natureza numa tecnologia aparentemente comum, seriam os assistentes domésticos criados por empresas como Google e Amazon. Para ensinar as crianças a serem “educadas”, e se utilizando do artifício de que a IA tem sentimentos e pode se ofender, há em um desses assistentes uma função que usa reforço positivo quando a criança demonstra ter aquilo que o sistema incorpora como sendo boas maneiras. Em contrapartida, se a criança não se comporta “bem” ela é orientada em direção a essa suposta educação/polidez. Ficam as perguntas: quem decide sobre o que é “polido” ou não? Isso incute nas crianças um respeito pelas entidades tecnológicas? Quais as consequências possíveis disso?

processos de aprendizagem de maneira a potencializá-los, não de reduzir as suas possibilidades. No decorrer da discussão a seguir, procuro avançar um pouco mais em alguns pontos.

2.5

Estendendo a discussão

Para buscar uma compreensão sobre o sentido de *aprendizagem* presente na *aprendizagem de máquina*, procurei olhar para as origens das máquinas usadas para ensinar e suas possíveis relações com as plataformas adaptativas de hoje. Essas máquinas geralmente não são consideradas como os primórdios da IA na educação; em vez disso, considera-se um recurso posterior, a Instrução Assistida por Computador (CAI-*Computer-Assisted Instruction*). Mas as máquinas analógicas dizem muito sobre as plataformas adaptativas hoje criadas para serem usadas na educação. Essa relação pode elucidar muito sobre concepções de aprendizagem vigentes no passado e ainda persistentes no presente.

Acontece que tecnologias de IA para apoiar a aprendizagem são criadas a partir da perspectiva que se tem da aprendizagem humana. Mas, em retorno, elas também podem ajudar a investigar como a aprendizagem de fato acontece. “Há várias razões pelas quais machine learning é importante. Claro (...), o desempenho da aprendizagem em máquinas pode nos ajudar a entender como animais e humanos aprendem” (Nilsson, 1996)⁴⁵. O trecho abaixo ilustra também esse aspecto:

AIED (Artificial Intelligence in Education) efetivamente envolve duas vertentes principais, complementares: o desenvolvimento de ferramentas baseadas em IA para dar suporte à aprendizagem e o uso desses recursos para ajudar a compreender a aprendizagem (como a aprendizagem acontece e outras questões que têm sido investigadas há bastante tempo pelas ciências da aprendizagem, e que podem ser aplicadas à sala de aula esteja sendo usada IA ou não). Por exemplo, modelando como estudantes fazem para resolver problemas de aritmética e identificando equívocos antes desconhecidos para educadores, pesquisadores e professores podem começar a entender mais sobre o processo de aprendizagem em si, que pode então ser aplicado às práticas convencionais de sala de aula. (Holmes, Bialik & Fadel, 2010, p. 83).⁴⁶

⁴⁵ No original: *There are several reasons why machine learning is important. Of course (...) the achievement of learning in machines might help us understand how animals and humans learn.* (Nilsson, 1996)

⁴⁶ No original: *AIED effectively involves two main complementary strands: developing AI-based tools to support learning, and using these tools to help understand learning (how learning happens and other questions that have long been investigated by the learning sciences, and which might be applied in classrooms whether or not AI is being used). For example, by modelling how students go*

Segundo Knox, Williamson e Bayne (2019, p. 8), a pesquisa em ML não tem equiparado a aprendizagem de máquina à aprendizagem humana. “Pesquisadores em machine learning são, certamente, conscientes das diferenças entre o reconhecimento de padrões e a inteligência humana”, dizem. Esses autores esclarecem que o que desejam é mostrar como ideias particulares sobre “aprendizagem”, especialmente aquelas envolvendo comportamento, têm atravessado a ciência da computação e as disciplinas educacionais por meio do uso de tecnologias de dados. No entanto, eu não diria – e nem a história indica – que essa visão esteja tão “segura”. É uma percepção que aparece entre cientistas de dados – o que se torna arriscado quando decisões relativas à educação possivelmente se concentram nas mãos de quem domina a técnica, mas não os fundamentos da educação ou da pedagogia⁴⁷. Um ponto importante, como indica Selwyn (2022), é que educadores tendem a ser permeáveis às ideias de uma IA generalista, da singularidade tecnológica e de alguns nichos tecnológicos que ainda são mais como delírios, como IAs se aproximando da senciência humana.

Um elemento crítico do discurso de que as máquinas aprendem como nós aprendemos é que, pela lógica, isso implicaria em que o contrário fosse igualmente verdadeiro, o que seria uma pressuposição problemática⁴⁸. Uma consequência disso seria a de que, equiparando-se a aprendizagem humana à capacidade de tecnologias educacionais cognitivas, se poderia acabar concluindo que essas tecnologias são capazes não só de aprender, mas de ensinar⁴⁹. Outra consequência relacionada seria

about solving an arithmetic problem and identifying misconceptions that might have been previously unknown to educators, researchers and teachers can begin to understand much more about the process of learning itself, which might then be applied to mainstream classroom practices. (Holmes, Bialik & Fadel, 2010, p. 83).

⁴⁷ Isso, somado a interesses que mais possam ter a ver mais com desenvolvimento econômico do que propriamente educativo, pode ter consequências importantes para a sociedade, como tem sido trabalhado por Williamson e colegas (2011, 2019), Selwyn (2019), O’Neil (2016), entre outros.

⁴⁸ Talvez a aprendizagem de máquina possa, sim, contribuir para a compreensão da aprendizagem humana, mas isso não seria por equivalência ou justaposição: pode-se caminhar para compreender a aprendizagem humana a partir da aprendizagem de máquina observando-se, por exemplo, os limites dos sistemas artificiais, ou o que acontece quando os utilizamos durante a nossa (constante) formação. Esta ideia está alinhada com o argumento principal da minha pesquisa, enquanto a primeira, de equivalência, não está.

⁴⁹ Selwyn (2019) desenvolve uma discussão sobre a hipótese da substituição de professores por robôs a partir não da dúvida sobre isso ser possível ou não, mas sobre ser desejável. Ele nos leva a refletir sobre o que queremos dos processos de ensino e aprendizagem, e para isso nos instiga a pensar sobre o que são e como ocorrem esses processos. Compreendê-los, afinal, é fundamental para que se estabeleça um debate não sobre a viabilidade, mas sobre qual sentido haveria em suplantiar os humanos por tecnologias nesses processos. Outro ponto importante levantado por Selwyn (2019, p. 16) refere-se ao fato de que toda tecnologia é humana; um robô professor seria uma combinação

uma possível transferência de responsabilidade a essas tecnologias. No trecho a seguir, Bannell (2017) exprime essa mesma preocupação:

Uma coisa é transferir tarefas cognitivas para máquinas ou facilitar a cognição distribuída. Outra é deixar, de um jeito ou de outro, a cargo de máquinas a direção do processo de aprendizagem. Esse é o caso da “aprendizagem automatizada” e da “aprendizagem adaptativa”. A “aprendizagem automatizada” envolve a construção de algoritmos que podem aprender a fazer previsões e alterar sua própria configuração à luz dessas previsões. Alguns autores falam do programa de computador que aprende com a experiência. Tal computador é uma máquina que pensa? Se o critério for desempenho operacional, então, pela aplicação do teste de Turing poderia se dizer que “sim”. Mas, se o critério for substancialmente cognitivo, a resposta não é tão simples. Como Searle ressaltou com seu famoso experimento mental do Quarto Chinês, saber manipular algoritmos e produzir resultados semelhantes aos de seres humanos não é equacionável a ser capaz de pensar. Isso deve-se ao pensamento requerer semântica, e, como Searle diria, a máquina “tem a sintaxe, mas não tem a semântica”. (Bannell, 2017, p. 35-36).

É uma preocupação presente entre filósofos que se interessam por inteligência artificial contribuir para a discussão sobre quais atividades caberiam exclusivamente a humanos e quais poderiam ser delegadas às máquinas ou aos sistemas artificiais. Smith (2019) acredita que definir o que seria *inteligência* é um aspecto central a essa discriminação: segundo o filósofo, não há dúvidas – sistemas artificiais irão transformar a existência humana; mas a compreensão quanto às suas capacidades, à sua confiabilidade, aos seus possíveis impactos e quanto às questões éticas emergentes demanda que entendamos o que a inteligência é, quais tipos de trabalho exigem quais tipos de capacidades e daí por diante. Longe de fazer isso de maneira utilitarista, Smith expõe esses aspectos para nos levar a refletir e analisar o nosso lugar no mundo, frente às transformações catalisadas pela IA, que traz novas reconfigurações ontológicas e epistemológicas; e propõe questões como: quem somos, para onde vamos, como queremos viver, no que acreditamos?

Smith, bem como outros filósofos cujo pensamento fundamenta teoricamente a tese aqui apresentada, como Ezequiel Di Paolo e Andy Clark, por razões sensivelmente diferentes não descartam a possibilidade de que se sintetize a cognição humana em sistemas artificiais. Há, porém, algo crucial que é comum ao pensamento desses filósofos: ao menos por enquanto, isto é, no estágio em que estamos quanto à IA, não estamos perto de conseguir reproduzir artificialmente o ser humano em sua totalidade. Em linhas gerais, pode-se dizer que o motivo para

de pessoas e máquinas, estruturas de códigos e o mundo material, os ambientes sociais. Ele sugere uma abordagem “sociotécnica” se quisermos atribuir um sentido à IA na educação.

isso é que, até o momento, apenas sistemas orgânicos contam com as características que seriam necessárias para haver atividade cognitiva. Portanto, ainda que, em tese, seja possível que sistemas artificiais adquiram tais características, tal feito ainda não aconteceu. Há vários pesquisadores que não aceitam isso, até por questões conceituais, e não empíricas. O que falta aos sistemas artificiais em relação às capacidades cognitivas humanas? E o que, em relação à forma como esses sistemas funcionam, nos revela sobre a nossa própria maneira de funcionar? Nos capítulos seguintes, avançarei em tópicos que são permeados por essas questões.

2.6

Conclusão do capítulo

Há anos a noção de inteligência humana e a concepção de cognição humana (por vezes consideradas sinônimos, ainda que não se tenha uma definição única de nenhuma das duas) vêm informando a inteligência artificial, e vice-versa. Como se vê pelas primeiras máquinas de ensinar, e pelo persistente interesse em ML na educação, hoje, a educação tem se inclinado à automatização do ensino, a partir daquilo que acredita que pode levar os alunos a “aprenderem melhor” (e mais rápido). Se a noção de inteligência, de cognição ou de aprendizagem é tomada pela metade, ou se o ponto de partida pressupõe noções de inteligência, cognição e aprendizagem que não contam a história completa, isso se reflete na IA – e na educação – também dessa maneira incompleta. O problema se intensifica quando essa concepção é estendida de uma forma que reduz a compreensão que se tem da aprendizagem humana.

Apesar de a realidade da inserção de ML na educação ainda ser incipiente, o que inclui o Brasil e Portugal, o que desejo é contribuir para que sejam discutidas as premissas do uso de ML como recurso para ensinar e aprender – até porque ML pode ainda não ser um recurso tão frequente nas salas de aula nesses países, mas já se encontra em diversos sistemas com os quais estamos em contato constante e que, desse modo, fazem parte do processo de formação humana (e afetam a sociedade de modo amplo). ML permeia as nossas vidas e o nosso processo contínuo de conhecer o mundo de várias maneiras e em vários formatos. Analisar pressupostos é importante porque, se por um lado o que se entende por aprendizagem em ML possivelmente não é o que se entende por aprendizagem na educação, por outro

lado a redução dessa compreensão poderá afetar a educação num futuro próximo, ou já está afetando. A colocação de Selwyn no trecho a seguir reforça isto:

Embora a IA atualmente continue sendo uma característica periférica para a maior parte das escolas e universidades, as maneiras como foram adotadas “precocemente” as ferramentas e tecnologias baseadas em IA já se tornaram uma ferramenta de educação e de práticas que suscitam uma série de questões que merecem ser levadas a sério nos próximos anos. (Selwyn, 2022, p. 2)

No próximo capítulo, aprofundo a investigação sobre o sentido de aprendizagem que tende a apoiar as tecnologias digitais, procurando olhar para o paradigma vigente nas teorias sobre a mente humana. Abordo os problemas que o cognitivismo enfrenta e que a inteligência artificial ajuda a expor, por meio da empiria. Da metade para o final do capítulo, são apresentadas as ideias de Hubert Dreyfus que considero condizentes com as concepções enativistas. Começo, assim, a demonstrar como um novo paradigma pode estar surgindo.

3

Tecnologias “cerebrais”

Tendo sido demonstradas, no capítulo anterior, as motivações subjacentes às tecnologias criadas para automatizar o ensino – e que ainda acompanham o discurso a promover o uso de tecnologias digitais envolvendo machine learning na educação – considero importante seguir com a análise das premissas que delineiam o desenvolvimento e dão suporte ao uso desses recursos. Com base em qual concepção de aprendizagem se argumenta ser bom, ou eficiente, ou valioso utilizar a aprendizagem de máquina nos processos de ensino e aprendizagem? Afinal, no que é que se fundamenta a promoção desses sistemas como tecnologias capazes de colaborar para educar ou para ensinar?

Neste capítulo, argumento que essa concepção é enraizada essencialmente na abordagem para a cognição humana que caracteriza a teoria computacional da mente, a qual privilegia a atividade cerebral e o processamento de informações. Essas são ideias que delineiam o paradigma cognitivista, o qual surgiu junto à revolução cognitiva na década de 1950. Tais concepções vêm dando lugar a outros focos de pesquisa na ciência cognitiva. Tendo como pontos nevrálgicos, entre outros, a simultaneidade entre percepção e ação e o corpo e a sua ação no mundo, abordagens contemporâneas procuram investigar a cognição com foco nas questões sobre como percebemos o mundo e como significamos ou damos sentido às coisas que percebemos.

Ao se debruçarem essencialmente sobre aquilo que está para além do processamento de informações – linha argumentativa típica do cognitivismo clássico – essas propostas mais recentes para a cognição, e que vêm ganhando corpo principalmente desde os anos de 1970 (Johnson, 2021), podem nos ajudar arepensar o sentido de aprendizagem que ainda predomina na educação. A esperada inovação trazida por tecnologias baseadas em ML corre o risco de estar assentada sobre pilares não tão sólidos, candidatos a serem postos em xeque na medida em que um novo paradigma ganha dimensão. Esta é a discussão deste capítulo.

3.1

Tecnologias do cérebro e para o cérebro

As tecnologias educacionais digitais baseadas em ML são, essencialmente, tecnologias “cerebrais” – no decorrer do capítulo, vou explicar por que razão as descrevo assim. Isso vale não exclusivamente para tecnologias desenvolvidas para fins educacionais, mas para as tecnologias baseadas em ML em geral. A história da IA tem um longo casamento com o pressuposto de que a inteligência é descorporificada, e esse pressuposto pode ser observado no trabalho da IA ao longo da história. Desde a abordagem simbólica para a IA até as redes neurais profundas, tem-se assumido que o cérebro e suas dinâmicas é o único responsável pela inteligência humana (Mitchell, 2021).

Ainda que não nomeie dessa maneira (isto é, como “cerebrais”) as tecnologias digitais, Williamson (2017) nos ajuda a enxergá-las assim. Sua perspectiva crítica para as tecnologias educacionais digitais explicita como o desenvolvimento desses recursos parte da perspectiva do cérebro e, por outro lado, como eles são também desenhados para desenvolver o próprio cérebro. Talvez isso não fosse um problema em si – tendo em conta que, por mais que se busque uma perspectiva para a cognição humana que vá além do cérebro, não se negue que esse órgão seja fundamental para a nossa atividade cognitiva, aprendizagem e percepção. No entanto, limitar a cognição ao cérebro pode se tornar uma pressuposição problemática quando a aprendizagem, concebida também de forma limitada ao desenvolvimento do cérebro, negligencia determinados fatores que vêm sendo apontados como igualmente essenciais para os processos que levam humanos a conhecer e fazer sentido do mundo. O problema parece mais grave quando o cérebro é concebido como um mero receptor de *inputs* que gera *outputs*, isto é, um órgão que seria basicamente um processador de informações. Vou procurar deixar mais claro por que considero isso um problema e a quais fatores exatamente eu me refiro como complementares ao cérebro na atividade cognitiva.

A noção de sujeito pensante que essencialmente acumula informações, processa e decodifica internamente aquilo que recebe do entorno tende a se alinhar à lógica geral subjacente ao desenvolvimento das tecnologias digitais baseadas em machine learning. Esse é o caso especialmente daquelas que são voltadas para a educação e se parecem, em termos de suas premissas, com máquinas de ensinar do

século passado, apesar de terem a particularidade de serem impulsionadas por *big data*. As tecnologias se voltam para o desenvolvimento do cérebro, o que encontra respaldo na neurociência – num movimento que vem se popularizando não apenas na educação, mas em diversas áreas da nossa sociedade, muito focada no cérebro. Williamson cita alguns exemplos nos trechos a seguir:

Os esforços para mapear o cérebro, para aproveitar as imensas bases de dados emergentes da neurociência atual, e em última análise para desenvolver o poder computacional para simular o funcionamento neural são os últimos indicadores de uma cultura em que o cérebro ocupa uma posição proeminente na sua capacidade de explicar todo o tipo de experiências humanas, decisões, capacidades, ações e relações. (Pykett, 2015, p. IX, apud Williamson, 2017, p. 189)⁵⁰

A neurociência se tornou o assunto de interesse apaixonado nos últimos anos. Fala-se agora de neuromarketing, neuroeconomia, neuropolítica, neurorrobótica, neurodireito, neuroeducação e neurotecnologia, e o prefixo "neuro" deu origem a críticas como neuromania, neuromitos e até mesmo neurostupididades. Todos esses domínios "neuro" somam-se a um "neurocomplexo" de afirmações opostas sobre o cérebro, um complexo que compreende o cérebro como "o orquestrador de nossos pensamentos, sentimentos, comportamentos e desejos, o objeto de pesquisa científica e intervenção, e um recurso cultural refinado e inesgotável e ponto de referência. O 'neuro' é, portanto, um campo complexo de verdades contestadas, reivindicações e contra reivindicações". (Williamson et al., 2011, pp. 145-146, apud Williamson, 2017, p. 191)⁵¹.

Nesse contexto, entusiastas das novas tecnologias, alimentadas por dados, tomam o cérebro como algo que pode ser medido, gerenciado e modificado, o que se transforma em combustível para o desenvolvimento de vários dispositivos tecnológicos.

Os defensores e desenvolvedores destas tecnologias inspiradas e estimuladas pelo cérebro encorajam os avanços atuais no campo da ciência de dados educacionais, tais como a análise da aprendizagem, aprendizagem adaptativa e plataformas de gestão das emoções, a concentrarem-se na neurologia humana e na base neurológica dos processos cognitivos. A nova variante neurocomputacional do imaginário do *big*

⁵⁰ No original: *Los esfuerzos por mapear el cerebro, aprovechar las inmensas bases de datos surgidas de la actual neurociencia, y al final desarrollar el poder informático para simular el funcionamiento neuronal, son los indicadores más recientes de una cultura en que el cerebro ocupa un lugar preeminente por su capacidad de explicar todo tipo de experiencias, decisiones, capacidades, actuaciones y relaciones humanas.* (Pykett, 2015, p. IX, apud Williamson, 2017, p. 189)

⁵¹ No original: *La neurociencia se ha convertido en objeto de apasionado interés en los últimos años. Hoy se habla de neuromarketing, neuroeconomía, neuropolítica, neurorrobótica, neuroderecho, neuroeducación y neurotecnología, y el prefijo "neuro" ha dado lugar a críticas como la neuromanía, los neuromitos y hasta las neurostupideces. Todos estos ámbitos "neuro" se suman al "neurocomplejo" de declaraciones opuestas sobre el cerebro, un complejo que entiende a este como "orquestrador de nuestros pensamientos, sentimientos, conductas y deseos, objeto de investigación e intervención científicas, y un refinado recurso cultural inagotable y punto de referencia. Lo 'neuro' es, pues, un complejo campo de disputadas verdades, afirmaciones y contraafirmaciones".* (Williamson et al., 2011, pp. 145-146, apud Williamson, 2017, p. 191)

data para a educação vê o próprio cérebro como o objeto de medição, gestão e modificação (Williamson, 2017, p. 190)⁵².

No parágrafo a seguir, é ilustrado um exemplo de dispositivos essencialmente “cerebrais” aplicados à educação:

Das neurociências, os primeiros resultados, voltados para aplicações educacionais, são tecnologias vestíveis que visam manter a atenção dos alunos direcionada para o professor, durante as aulas. Tiaras utilizadas por alunos possuem sensores que recebem sinais do cérebro, do tipo Eletro Encefalograma (EEG), e também emitem sinais ao cérebro, para monitorar a atenção do aluno. Os equipamentos vestíveis, em geral, estão conectados à internet e apontam um dos caminhos para a Internet das Coisas (Internet of Things - IoT), em aplicações educacionais. (Vicari, 2021, p. 79-80)

Williamson (2017, p. 192) assinala que, cada vez mais, os estudos sobre o cérebro têm sido associados às dinâmicas sociais, dando lugar à aparição de novos tipos de estudos “biossociais” em cuja base estão “a permeabilidade do corpo e do cérebro ao seu entorno social, material e técnico”. Segundo o autor, a ideia é reparar a divisão feita artificialmente entre as ciências sociais e as da vida, uma vez que a sociedade foi apartada da natureza de uma maneira que deveria ser superada a partir de uma investigação interdisciplinar. Conceitos que surgem dessa mescla entre as ciências sociais e biológicas seriam, por exemplo, “neuroplasticidade” e “epigenética”:

Em termos simples, a neuroplasticidade reconhece que o cérebro se adapta constantemente a estímulos externos e ambientes sociais, enquanto a epigenética reconhece que a experiência social modula o corpo ao nível genético (Chung et al, 2016; Tose e Abi-Rached, 2013, apud Williamson, 2017, p. 192)⁵³.

Ainda segundo Williamson, a união dessas áreas constitui uma potente fonte de dados, usados essencialmente para alimentar as próprias tecnologias basicamente inspiradas no cérebro e desenvolvidas para o cérebro. Conclui-se que

⁵² No original: *Defensores y desarrolladores de estas tecnologías inspiradas en el cerebro y de estímulo cerebral fomentan los actuales avances en el campo de la ciencia de datos educativos, como la analítica del aprendizaje, las plataformas de aprendizaje adaptativo y de gestión de las emociones, para centrarse en la neurología humana y la base neurológica de los procesos cognitivos. La nueva variante neurocomputacional del imaginario de los big data para la educación ve en el propio cerebro el objeto de medición, gestión y modificación.* (Williamson, 2017, p. 190)

⁵³ No original: *Dicho de forma sencilla, la neuroplasticidad reconoce que el cerebro se adapta constantemente a los estímulos externos y los entornos sociales, mientras que la epigenética admite que la experiencia social modula el cuerpo a nivel genético.* (Chung y col, 2016; Tose y Abi-Rached, 2013).

emerge, assim, uma indústria que se retroalimenta, como o trecho a seguir deixa a entender.

Um aspecto fundamental dos estudos biosociais críticos é a análise das tecnologias envolvidas na produção de neurociências, das organizações e outros agentes que as promovem e divulgam, e dos setores em que são utilizadas. A própria neurociência tornou-se uma grande empresa de dados através de imagens de alta resolução do cérebro e projetos de simulação, por exemplo, o Projeto Cérebro Humano⁵⁴. Estes projetos, por vezes classificados como neurociência, geram torrentes de grandes dados que são utilizados para transformar formas de compreender o cérebro como um órgão neuromorfológico multiconectado, dinâmico e socialmente adaptável, e para impulsionar o desenvolvimento de tecnologias e processos computacionais baseados no cérebro (Nature Neuroscience, 2014). A neurocomputação é um campo fundamental em que a plasticidade e a natureza epigenética do cérebro é utilizada para ligar a perícia neurocientífica ao desenvolvimento técnico, interesses comerciais e objetivos governamentais. (Williamson, 2017, p. 195)⁵⁵

Com o *big data*, as imensas massas de dados de alunos que têm nas mãos e que são geradas pelas suas próprias invenções, as empresas desenvolvedoras de tecnologias que procuram definir como educacionais podem gerar ainda mais dados. Um ponto sensível disso é o risco de resumir os alunos, seu “perfil”, seu “comportamento” e seu desempenho (tais como compreendidos no contexto) a aquilo que fica refletido em seus dados. No entanto, se defende que é assim que se pode oferecer uma experiência de aprendizagem personalizada, resolvendo os pontos fracos de cada um, medindo as respostas neuronais em tempo real e, em última instância, melhorando a “cognição” a partir daquilo que é revelado sobre a progressão da aprendizagem a partir dos dados coletados. “Entraram em cena pressupostos controversos sobre a aprendizagem humana como um processo computacional ou como um processo determinado pela variância genética, todos argumentados a partir da ideia de personalização” (Williamson, 2017, p. 213).⁵⁶

⁵⁴ The Human Brain Project (<https://www.humanbrainproject.eu/en/>).

⁵⁵ No original: *Un aspecto fundamental de los estudios biosociales críticos es el análisis de tecnologías que intervienen en la producción de la neurociencia, de las organizaciones y otros agentes que las promueven y difunden, y de los sectores en los que se utilizan. La propia neurociencia se ha convertido en una empresa de macrodatos a través de proyectos de imágenes y simulación cerebrales de alta resolución, por ejemplo, el Proyecto Cerebro Humano. Estos proyectos, a veces clasificados como neurociencia, generan torrentes de big data que se utilizan para transformar los modos de entender el cerebro como órgano de múltiples conexiones, neuromorfología dinámica y socialmente adaptativa, y para impulsar el desarrollo de tecnologías y procesos informáticos basados en el cerebro (Nature Neuroscience, 2014). La neurocomputación es un campo fundamental en el que la plasticidad y la naturaleza epigenética del cerebro se utilizan para conectar la experiencia neurocientífica con el desarrollo técnico, los intereses comerciales y los objetivos de los gobiernos. (Williamson, 2017, p. 195)*

⁵⁶ No original: *Han entrado en escena polémicos supuestos sobre el aprendizaje humano como proceso informático o como proceso determinado por la varianza genética, todos argumentados con la idea de personalización. (Williamson, 2017, p. 213)*

Controversos desdobramentos possíveis, essencialmente sociais e éticos, decorrem da força crescente de uma indústria do design instrucional baseada em *big data* e que preconiza a união entre fatores biológicos e sociais.

Desenvolvedores de tecnologias ditas *educacionais* não escondem que criam seus sistemas com base no cérebro humano e para melhorá-lo; por isso reafirmo que são, afinal, tecnologias *do cérebro e para o cérebro*⁵⁷. Mais precisamente, e também talvez com viés mais problemático, partem de um cérebro humano tomado estritamente como um consumidor e concomitante gerador de dados. Defendem que, assim, estão produzindo modelos inteligentes de ensino e aprendizagem porque são capazes de criar sistemas que geram e leem esses dados de uma maneira que potencializa a qualidade da aprendizagem. Eles chegam a proclamar que a IA aplicada à educação – nomeadamente na forma de machine learning alimentado por *big data*, recursos classificados comumente como tecnologias cognitivas – representa um novo *paradigma* para esse campo.

Até recentemente, a computação era programável, baseada em entradas, instruções (código) e saídas definidas pelo homem. Os sistemas cognitivos estão num **paradigma** completamente diferente de sistemas que compreendem, raciocinam e aprendem - sistemas, em suma, que pensam. O que poderá isto significar para os educadores? Acreditamos que os sistemas cognitivos podem alargar as capacidades dos educadores, fornecendo-lhes conhecimentos profundos e ajuda especializada através do fornecimento de informação de uma forma atempada, natural e utilizável. Estes sistemas cumprirão o papel do assistente, um complemento, e não um substituto, da arte e do ofício do ensino. No cerne dos sistemas cognitivos estão as capacidades analíticas avançadas. Em particular, os sistemas cognitivos visam responder às perguntas: "O que vai acontecer?" e "O que devo fazer?" (King et al., 2016, p. 9, apud Williamson, 2017, p. 205, grifo meu)⁵⁸.

⁵⁷ Em Williamson, 2017, p. 207: *É verdade que os sistemas cognitivos são máquinas inspiradas pelo cérebro humano. Mas também é verdade que estas máquinas irão inspirar o cérebro humano, dar-nos uma maior capacidade de raciocínio e refazer a forma como aprendemos.* (Kelly, 2015, p. 11)/ No original: *Es verdad que los sistemas cognitivos son máquinas inspiradas en el cerebro humano. Pero también lo es que estas máquinas inspirarán al cerebro humano, nos darán mayor capacidad de razonamiento y recablearán nuestra forma de aprender.* (Kelly, 2015, p. 11)

⁵⁸ No original: *Hasta hace poco, la computación era programable, estaba basada en inputs, instrucciones (código) y outputs de definición humana. Los sistemas cognitivos están en un paradigma completamente distinto de sistemas que entienden, razonan y aprenden. Sistemas, en resumen, que piensan. ¿Qué puede significar esto para los educadores? Consideramos que los sistemas cognitivos pueden ampliar las capacidades de los educadores aportándoles profundos conocimientos sobre este ámbito y ayuda especializada a través de la provisión de información de forma puntual, natural y utilizable. Estos sistemas cumplirán la función del asistente, una función complementaria, no sustitutoria, del arte y el oficio docentes. En el centro de los sistemas cognitivos están unas avanzadas capacidades analíticas. En particular, el objetivo de los sistemas cognitivos es responder las preguntas: "¿Qué ocurrirá?" y "¿Qué debería hacer?".* (King y col., 2016, pág. 9)

Cabe questionar se há, de fato, uma mudança de paradigma capitaneada por esses sistemas, ditos cognitivos, e que teoricamente “pensam”. Inovação não deve, afinal, ser confundida com mudança de paradigma. Deve-se atentar para o fato de que a concepção de cognição que apoia esses sistemas ainda se mostra assentada sobre o cérebro como eixo central; assim, fica restrita a um suposto processador central ao qual a atividade cognitiva (e a mente) se submeteria por completo. Por outro lado, quando olhamos para a ciência cognitiva e os seus desenvolvimentos, especialmente nos últimos 30 anos – intimamente entrelaçados com os avanços nas pesquisas em filosofia da mente, robótica e inteligência artificial – podemos observar que há, sim, no campo das pesquisas em cognição (unindo diversas disciplinas e programas de pesquisa) um processo rumo à superação de um determinado paradigma. Trata-se de uma transformação ainda em curso, que vem se desenvolvendo de maneira lenta, profunda e enraizada em inúmeros fatores que adquirem força passo a passo. Isso é natural, já que mudanças de paradigma não são rápidas, até por serem estruturais.

Essa movimentação nas ciências cognitivas gera, de fato, novos terrenos; não se limita a reorganizar o que já estava apoiado sobre os mesmos pilares. E, nesse caso, se trata justamente de uma revisão da concepção do cérebro como um processador central, um consumidor de dados (*inputs*) do ambiente, gerador de novos dados (*outputs*); da revisão de uma concepção de cognição equiparada ao cérebro que remete a Descartes e Kant, pois se mostra dualista, separando mente/cérebro e corpo; e internalista, concentrando no interior do indivíduo (o sujeito “desprendido” ou “desengajado”; ver Bannell et. al, 2016; Taylor, 2005) as suas capacidades cognitivas, que utilizariam o meio externo apenas como fonte de estímulos e informações. Caminha-se para buscar compreender cada vez melhor, e com mais detalhes, como a atividade cerebral se conecta com a atividade do corpo como um todo, gerando cognição; há uma tentativa de olhar para elementos variados do ambiente como essenciais para explicar a cognição. Abordagens contemporâneas procuram investigar a cognição com foco nas questões sobre como percebemos o mundo, como significamos ou damos sentido às coisas que percebemos. Mesmo para tentar entender o que acontece individualmente num ser cognoscente, essas abordagens valorizam a intersubjetividade. Também na educação já se concebe há muito o desenvolvimento cognitivo como algo que acontece no domínio relacional, não apenas individual.

Sendo assim, tecnologias que abordam a neurociência a partir das conexões entre processos sociais, culturais, biológicos e neurológicos podem representar um avanço em relação a outras que se restringem ao cognitivismo clássico que isolava a mente dessas dimensões. Mesmo para fazer justiça à neurociência enquanto área de pesquisa, é importante ressaltar que sua investigação pode ter múltiplas apropriações. Mitchell (2021, p. 7) classifica a neurociência como uma das disciplinas a partir das quais a cognição corporificada floresceu, e ressalta a área como detentora de um papel de relevância no crescimento das abordagens da cognição corporificada. Segundo a autora, trata-se de um campo que mostra que as estruturas neurais cognitivas “são ricamente ligadas a aquelas que controlam os sistemas sensoriais e motores, e que o pensamento abstrato explora mapas neurais baseados no corpo”. Citando o trabalho sobre metáforas de Lakoff e Johnson, ela afirma que

Resultados vindos da psicologia cognitiva e da linguística indicam que muito, se não todos, os nossos conceitos abstratos são enraizados em modelos internos físicos, baseados no corpo, revelados em parte pelos sistemas físicos de metáforas encontrados na linguagem corriqueira. (Mitchell, 2021, p. 7)

Por outro lado, enquanto o cérebro seguir como sinônimo de *sistema cognitivo* – a partir do qual, em última medida, o que se quer é observar, ler e interpretar esse órgão e suas conexões – não se poderia proclamar uma quebra de paradigma. Afinal, ter em vista que fatores corporais, ambientais e biológicos influenciam a atividade do cérebro em alguma medida é tomar esses fatores como algo que *impacta*, mas não necessariamente *constitui* a mente. Algo sensivelmente diferente é revisar e (re)equalizar o peso das relações entre o cérebro e esses outros elementos, admitindo inclusive que é dessas relações que emerge a mente. Ainda mais radical, nesse sentido, é a proposta de condicionar a própria existência da mente, e da cognição, ao corpo e à sua atividade no mundo. Mesmo quando se trata de neurociência em si, enquanto área de pesquisa, é importante lembrar que ela pode crescer em harmonia com essa crescente perspectiva da mente corporificada, como lembra Johnson no trecho a seguir:

Em meados da década de 1970, o cenário estava começando a mudar, na medida em que vários programas de pesquisa em filosofia, psicologia, linguística, ciência da computação e **neurociência** foram se unindo para formar uma nova classe, uma segunda geração da ciência cognitiva, que reconheceu como nossa corporificação é importante para tudo o que experienciamos, pensamos, significamos, comunicamos e fazemos. Quanto mais os pesquisadores estudavam como as pessoas de fato

pensam – de maneira oposta a como filósofos, lógicos e cientistas da computação diziam que deveriam pensar – mais claro ficava que nossos corpos impõem as próprias condições da nossa experiência, do nosso pensamento, e da nossa comunicação com outras pessoas. Essa nova perspectiva veio a se tornar conhecida como *a ciência cognitiva da mente corporificada* (*cognitive science of the embodied mind*) ou, mais simples, *ciência cognitiva corporificada* (*embodied cognitive science*) (Lakoff e Johnson, 1999; Gibbs 2006; Feldmann, 2006). (Johnson, 2021, p. 107, grifo meu).

Mas o que poderia, então, caracterizar uma mudança de paradigma, de fato, nesse contexto? Como podemos saber se há ou não uma transformação dessa natureza em curso na ciência cognitiva? Os proponentes do enativismo ressaltam que uma mudança de paradigma provoca o deslocamento de um eixo do qual muitas teorias partem. Stewart, Gapenne e Di Paolo (2010, introdução⁵⁹) defendem que a enação/enact é um paradigma recente capaz de renovar a ciência cognitiva como um todo; Stewart (2010) afirma que se trata de um novo paradigma para a ciência cognitiva porque cumpre os requisitos básicos para isso, os quais seriam:

I. Oferecer uma proposta genuína de resolução do problema mente-corpo;

II. Oferecer uma articulação genuína entre uma série de disciplinas – pelo menos entre a psicologia, a linguística e a neurociência.

Stewart (2010) ressalta, ainda, que a ciência cognitiva deve sua própria existência à Teoria Computacional da Mente, a qual – por mais que tenha seus “defeitos” – sem dúvida preenche esses requisitos (ainda que seja à sua maneira, e na sua lógica, é claro). Essa teoria paradigmática, surgida com a revolução cognitiva que eclodiu nos anos de 1950, ainda é a mais aceita e subjaz à grande maioria das pesquisas que envolvem o cérebro e a cognição, equivalendo a mente ao cérebro e enfatizando o caráter de processador(a) de informações; não se trata, portanto, de premissas presentes apenas na educação ou nas tecnologias educacionais.

Graças à visão computacional-representacional da mente, a ciência cognitiva pôde definir-se como o programa científico de especificar a arquitetura funcional da inteligência, construindo novos formalismos e derivando suas implicações específicas para serem testados em última instância contra evidências empíricas e pela construção na forma de maquinaria artificialmente inteligente. Em suma, graças à visão computacional-representacional da mente, foi possível que a ciência avançasse. Isso é o que marca a visão computacional-representacional da mente como um paradigma científico frutífero, não sua solidez teórica ou sua incapacidade de lidar com problemas que os cientistas cognitivos escolhem ignorar (quando, discutivelmente, não deveriam). Aqueles de nós que sustentam que essa perspectiva é falha ainda devem reconhecer sua fecundidade e ser realistas sobre esse fato

⁵⁹ A primeira página da introdução do livro, onde encontrei as informações presentes neste parágrafo, não é numerada.

frequentemente negligenciado: o sucesso de um paradigma científico não é julgado apenas no tribunal da consistência lógica e da evidência empírica. A sua maturidade, por definição, reside também no facto de fornecer o conjunto certo de ideias e ferramentas para fazer face aos problemas que se coloca, independentemente da postura crítica que aponta para os seus pontos cegos como sendo também problemas genuínos e relevantes. (Di Paolo, 2009, não paginado)⁶⁰

Cabe atermo-nos por um instante à questão dos possíveis pontos fracos ou pontos cegos de um paradigma, e qual seria o momento em que esses problemas passam a fazer a diferença a ponto de pôr o paradigma em xeque. O conceito de *anomalia*, apontado por Chalmers a respeito da teoria de Kuhn para um paradigma, ajuda-nos a pensar nesse sentido.

A mera existência de enigmas não resolvidos dentro de um paradigma não constitui uma crise. Kuhn reconhece que os paradigmas sempre encontrarão dificuldades. Anomalias haverá sempre. É somente sob conjuntos especiais de condições que as anomalias podem se desenvolver de maneira a solapar a confiança num paradigma. Uma anomalia será considerada particularmente séria se for vista atacando os fundamentos de um paradigma e resistindo, entretanto, persistentemente, às tentativas dos membros de uma comunidade científica normal para removê-la. (...) Quando as anomalias passam a apresentar problemas sérios para um paradigma, um período de acentuada insegurança profissional começa. (Chalmers, 1993, p. 117)

Nas críticas ao cognitivismo, a robótica e a IA têm um papel fundamental. Afinal, a pesquisa empírica nesses campos contribui para tornar visíveis os problemas do programa cognitivista, revelando especialmente as lacunas que surgem quando se negligencia a importância do corpo e sua relação com o ambiente quando se deseja compreender a mente humana. Nas seções seguintes, irei abordar o entrelaçamento da inteligência artificial com a ciência cognitiva, apontando algumas das questões que contribuíram para tangibilizar aspectos sensíveis da mente humana.

⁶⁰ No original: *Thanks to the computational-representational view of the mind cognitive science can define itself as the scientific programme of specifying the functional architecture of intelligence, building new formalisms and deriving their specific implications to be ultimately tested against empirical evidence and by construction in the form of artificially intelligent machinery. In short, thanks to the computational-representational view of the mind, it has been possible for science to move on. This is what marks the computational-representational view of the mind as a fruitful scientific paradigm, not its theoretical soundness or its in/ability to deal with problems that cognitive scientists choose to ignore (when they arguably should not). Those of us who sustain that this perspective is flawed must still recognise its fruitfulness and be realistic about this oft-neglected fact: the success of a scientific paradigm is not solely judged at the court of logical consistency and empirical evidence. Its maturity, by definition, also lies in the fact that it provides the right set of ideas and tools to tackle the problems that it sets itself, never mind the critical stance that points to its blindspots as also being genuine and relevant problems.* (Di Paolo, 2009, não paginado)

3.2

A eclosão do cognitivismo e da inteligência artificial

Quando foram criadas as primeiras máquinas de ensinar para serem usadas em sala de aula, a ciência cognitiva ainda não havia se constituído como área de pesquisa; as máquinas de Skinner emergem em meio ao behaviorismo. Com a eclosão da chamada revolução cognitiva nos anos de 1950, esse paradigma passaria a ceder espaço para o paradigma cognitivista. Como explicam Di Paolo, Buhrmann e Barandiaran (2017, p. 14), a pressuposição predominante no programa de pesquisa cognitivista que emergiu então era a de que os processos cognitivos seriam fundamentalmente como programas de computador: a mente estaria para o cérebro da mesma forma que um software estaria para o hardware em que ele roda. Ainda segundo os autores (2017, p. 14), à época as investigações de modo geral centraram-se na cognição humana abstrata – elementos como o raciocínio conceitual, a resolução de problemas, a produção e a compreensão de linguagem, entre outros. Estados mentais – desejos, crenças, percepções etc. – eram concebidos como estados do cérebro, podendo ser identificados pelo seu papel funcional; por exemplo, suas relações com outros estados mentais e o seu papel na produção de um determinado comportamento.

O estudo do cérebro foi concebido, nesse contexto, em termos das computações que poderiam nele ser instanciadas. Um ponto fundamental também destacado por Di Paolo, Buhrmann e Barandiaran (2017, p. 14-15) é que, a essa altura da pesquisa em cognição, a atenção geral estava voltada não para a maneira como ocorre a construção de significado, mas para o processamento de informações. Resulta disso que um problema central da proposta cognitivista, na visão desses autores – proponentes do enativismo – é a sua ênfase nas manipulações sintáticas descorporificadas, que não endereçam a questão semântica sobre como qualquer tipo de significado poderia emergir de tais cálculos. Esse problema aparece na experiência do Quarto Chinês proposta por Searle (1980) e foi alvo da atenção do filósofo Hubert Dreyfus (1972, 1992), o qual durante toda a sua carreira se debruçou sobre essas questões.

A inteligência artificial emergiu, como área de pesquisa, praticamente junto com a eclosão da ciência cognitiva enquanto campo. Data exatamente de 1950 o célebre artigo “Computing Machinery and Intelligence”, de Alan Turing. Em 1956

surgiu o termo “artificial intelligence” – cunhado como parte de um projeto e de um workshop ocorrido na Dartmouth College, nos EUA. Não há, até hoje, uma definição única do que significa inteligência artificial – tampouco uma definição única de inteligência. Para Nilsson,

(...) a inteligência artificial é aquela atividade dedicada a criar máquinas inteligentes, e inteligência é aquela qualidade que habilita uma entidade a funcionar apropriadamente e antecipando o seu ambiente. De acordo com essa definição, muitas coisas – humanos, animais, e algumas máquinas – são inteligentes” (Nilsson, 2010, p. xviii).

A gestação da IA teve início, porém, antes do evento da Dartmouth College. Remete ao surgimento da cibernética e da ciência da computação, nos anos de 1940 (Williamson e Eynon, 2020). Um simpósio ocorrido em 1948 no California Institute of Technology, em Pasadena, marcou de forma definitiva esse então emergente campo; a conferência, de caráter interdisciplinar, tinha como focos os tópicos do controle do comportamento pelo sistema nervoso e as maneiras como o cérebro poderia ser comparado a um computador. Na ocasião, o psicólogo Karl Lashley ministrou uma palestra que, como coloca Nilsson (2010, p. 49), contribuiu para pavimentar o caminho para a ciência cognitiva como viemos a conhecê-la. Em sua apresentação, Lashley responsabilizou o behaviorismo pela “visão estática do funcionamento cerebral” e deixou claro que, se quisessem explicar as habilidades humanas de planejamento e linguagem, os psicólogos precisariam “começar considerando estruturas dinâmicas e hierárquicas”.

Williamson e Eynon (2020) ressaltam que, entre os anos de 1960 e 1990, a IA foi caracterizada por pesquisas focadas na codificação dos princípios do raciocínio humano, sendo permeada por tentativas de simular a inteligência humana e, mais tarde, por “expert systems” que emulavam os processos de tomada de decisão de especialistas a partir de bases de dados. Como conta Nilsson (2010, p. 47), diversos experimentos foram conduzidos naquele período inicial da IA, como “resolver quebra-cabeças, jogar xadrez e damas, provar teoremas, responder a questões simples e classificar imagens”. Apesar de alguns desses projetos corresponderem a tentativas condicionadas a laboratórios, outros mais diretamente relacionados ao mundo real e ao cotidiano também foram conduzidos – como, por exemplo, as investidas na automação da leitura e da tradução de idiomas. De todo modo, o autor ressalta que um importante elemento a fazer com que pesquisas na

IA e na psicologia cognitiva se tornassem tão entrelaçadas foi o fato de, em meio à efervescência que marcou a primeira fase da IA⁶¹, alguns pesquisadores estarem interessados sobretudo na maneira como o cérebro humano funciona e resolve problemas. Foi a partir dessa perspectiva que eles orientaram as suas pesquisas em busca de maneiras de levar máquinas a fazerem as mesmas atividades tipicamente cerebrais.

Essa premissa, por assim dizer, *cerebral* para a IA está presente tanto no cognitivismo como no conexionismo. No cognitivismo clássico, predominavam os sistemas simbólicos lineares, enquanto o conexionismo tem como característica o processamento paralelo distribuído (Clark, 2014)⁶². Uma das principais características das redes neurais artificiais são as conexões simultâneas e multidirecionais entre as unidades (os neurônios artificiais) gerando resultados diferentes. Técnicas de machine learning “são implementadas em arquiteturas geralmente conhecidas como *neural networks* (redes neurais), por conta de sua similaridade topológica com a maneira como o cérebro é organizado no nível neural” (Smith, 2019, p. 47)⁶³.

É essa capacidade crescente de processamento simultâneo de informações que potencializa as máquinas de IA de maneira a provocar uma reorientação na história do campo. Por conta do incremento de tal capacidade, os sistemas tornam-se habilitados a lidar com muito mais dados, em diversas camadas de profundidade (em contrapartida, a própria sofisticação da capacidade de processamento de dados desses sistemas só se torna possível porque há cada vez mais facilidade na produção desses dados). Como lembra Smith (2019)⁶⁴, o computador de um milhão de dólares, típico da IA em sua primeira fase ou GOFAI (*Good Old-Fashioned AI*), ocupava uma sala inteira e contava com menos de um milionésimo do poder de

⁶¹ Apesar da empolgação que marcou os anos iniciais da IA, essa fase também foi pontuada por períodos de “inverno”, conhecidos como “AI Winter” (Williamson & Eynon, 2020).

⁶² Em minha dissertação de mestrado (2019), explico em detalhes o cognitivismo e o conexionismo.

⁶³ No original: *These techniques are implemented in architectures often known as ‘neural networks’, because of their topological similarity to the way the brain is organized at the neural level* (Smith, 2019, p. 47).

⁶⁴ Smith (2019) usa a expressão “inebriada”, ou heady, em inglês, para se referir à forma como a imprensa se encontra em relação aos avanços tecnológicos no mundo da IA. Os meios de comunicação de massa, sites e blogs de tecnologia são alguns dos principais responsáveis pela disseminação das novidades – também sobre IA e machine learning. Ao se encontrarem intoxicados por uma espécie de otimismo generalizado, certamente levam a público essas suas impressões. Por outro lado, há críticos bastante preocupados com os usos do machine learning no contexto educacional, como Williamson (2017), Watters (2021), O’Neil (2016) e Selwyn (2019).

processamento dos telefones celulares atuais. Ressaltando que nós ainda não temos uma definição daquilo que significa *pensar* – a despeito de toda a evolução tecnológica em curso – o filósofo assume que:

Usando ideias diferentes, massas de dados coletados, e hardware radicalmente melhorado, os resultados do machine learning são realmente impressionantes” (...) “Novas conquistas são publicadas todos os dias” (...) Certamente os pesquisadores de IA estão mais animados e otimistas do que estiveram em 50 anos; não é somente a imprensa que está inebriada. Eu também admito que os desenvolvimentos predizem mudanças profundas na natureza da sociedade e no nosso entendimento de nós mesmos (Smith, 2019, p. 54)⁶⁵.

No cognitivismo e no conexionismo, percebe-se claramente que aquilo que importa para a cognição está dentro da cabeça, enquanto, no espaço externo, encontra-se aquilo que o cérebro precisa receber, absorver, para poder operar, enfim, processando dados. Pode-se dizer que se trata de uma relação em que o interno é mais importante que o externo, pois todo o trabalho de realizar essa digestão de informações cabe ao interior do sujeito cognoscente. Se o sujeito pensa, ele existe; para pensar, só é preciso que haja mente, a qual estaria apartada do corpo. O mundo é um fornecedor de dados que a mente (o cérebro) precisa absorver e decodificar.

Por conta disso, além do caráter computacional, a mente tal como definida por uma premissa cognitivista, demanda a existência de representações mentais para que seja realizado o contato do ser cognoscente com o mundo. Conseguiríamos, assim, ter acesso não ao mundo em si, mas a representações dele, como se fossem espelhos daquilo que de fato existe no mundo⁶⁶. O debate sobre a necessidade ou não dessas mediações para a atividade cognitiva, e até mesmo sobre o que seriam exatamente as representações, é bastante extenso e não é o foco central da investigação desta tese. Apesar disso, é um tema de certo modo incontornável quando se analisa a cognição e especialmente quando se busca uma perspectiva corporificada para a mente; afinal, admitindo-se que a ação corporal no ambiente é o cerne da atividade cognitiva, isso demanda que se revise o sentido das

⁶⁵ No original: *Using different ideas, masses of collected data, and radically improved hardware, the results of machine learning are genuinely impressive* (...) “*New accomplishments are published almost daily*” (...) “*Certainly AI researchers are more excited and optimistic than they have been in 50 years; it is not just the press that is heady. I too agree that the developments portend profound changes to the nature of society and our self-understanding* (Smith, 2019, p. 54).

⁶⁶ Ver Leporace et al. (2019)

representações nesse contexto⁶⁷. Passarei por esse assunto novamente quando isso se fizer necessário.

3.3

Problemas do paradigma cognitivista evidenciados pela IA

Quando a inteligência artificial começou a se desenvolver, era visível a confiança dos programadores quanto à possibilidade de reproduzir artificialmente as capacidades humanas. Centrados na ideia de que a mente era o cérebro, e de que o cérebro seria como um computador, acreditavam basicamente que poderíamos fazer emergir uma mente a partir de um computador artificial. “O problema com a metáfora da mente como máquina é que nem a mente humana e nem o cérebro humano é como computadores digitais” (Rohde, 2010, p. 11). A autora explica que um computador digital é um dispositivo que mapeia símbolos obtidos como *inputs* e gera símbolos na forma de *outputs*, seguindo, para isso, regras sintáticas; sendo assim, ainda que humanos sejam bons em fazer esses mapeamentos – mais do que a maioria dos outros animais – isso não é tudo o que fazemos.

A proposta da chamada *Good Old Fashioned AI* ou *GOF AI* constituiu a era da IA simbólica, marcada pela *Physical Symbol System Hypothesis* (PSSH) de Allen Newell e Herbert Simon, que apareceu no final dos anos de 1950 (Dreyfus, 1992; Mitchell, 2021). Por mais que, após as concepções simbólicas para a IA, tenham surgido no campo outras vertentes, inspiradas em estatística e na arquitetura neural cerebral, elas permanecem cérebro centradas, mantendo o corpo em segundo plano.

As redes neurais profundas de hoje são parecidas com o proverbial cérebro-no-tanque: absorvendo passivamente dados do mundo e fornecendo instruções de comportamento sem interagir ativamente no mundo com qualquer tipo de corpo. Evidentemente, robôs e veículos autônomos são diferentes por terem uma presença física no mundo, mas até o momento os tipos de interações físicas que eles têm, e o feedback para sua “inteligência”, é bastante limitado⁶⁸. (Mitchell, 2021, p. 6)

⁶⁷ Ver Rolla (2021, p. 39-40), que especifica dois tipos de argumentos contra a tese de *que toda cognição envolve representações*, um de caráter ontológico e outro de caráter epistemológico. Tendo nascido junto à psicologia ecológica de James Gibson, o argumento epistemológico promove a ideia de que não é necessário invocar representações mentais para explicar todo modo de cognição. Já o argumento ontológico – mais ambicioso, nas palavras do autor – é postulado por Hutto e Myin (2013) e defende que representações mentais não existem.

⁶⁸ No original: *Today's deep neural networks are akin to the proverbial brain-in-a-vat: passively taking in data from the world and outputting instructions for behavior without actively interacting*

De todo modo, as aplicações práticas no campo da inteligência artificial ajudaram a demonstrar como a ideia da mente computacional deixa de fora aspectos da interação entre corpo e ambiente que podem ser centrais à compreensão da cognição.

Ao pôr à prova os princípios cognitivistas no projeto de aplicações inteligentes, interfaces de usuário ou robôs eficientes que devem lidar com o mundo real, tornou-se evidente que era necessário prestar mais atenção ao corpo em sua interação física direta com o meio. (Di Paolo, Buhrmann & Barandiaran, 2017, p. 15)⁶⁹.

Ainda conforme esclarecem Di Paolo et al. (2017), concebidos a partir da ideia linear de percepção-ação – que, no caso dos robôs, se traduzia num fluxo desde a informação sensorial até a ativação dos efetores⁷⁰ – essa linearidade mostrou-se um obstáculo para o desenvolvimento de robôs capazes de interações com o ambiente consideradas “simples”, mas que envolvem diversos elementos simultâneos e escalas de tempo diversas, concomitantes. O ponto é que esses ambientes caóticos, com diversos fatores ocorrendo simultaneamente, não são a exceção, mas a regra. Uma das soluções que passou a ser buscada por roboticistas foi a construção de robôs com diversas camadas incrementais de controle.

Por conta da separação entre símbolos e significados, nas últimas décadas diversos problemas originaram-se da metáfora da mente como máquina: o *frame problem* (tratarei dele mais adiante); o problema do crédito pelo sucesso em machine learning (quando se segue diversos passos paralelos para resolver um problema complexo, como saber qual deles foi responsável pelo sucesso?); o problema filosófico de atribuir significado a símbolos ou como símbolos adquirem seu significado; o problema de compreender como aspectos que são processados em diferentes canais ou partes do cérebro podem ser reunidos de modo a formar uma percepção coerente do mundo, entre outros (ver Rohde, 2010, p. 11-12).

in the world with any kind of body. Of course, robots and autonomous vehicles are different in that they have a physical presence in the world, but to date the kinds of physical interactions they have, and the feedback to their “intelligence” is quite limited. (Mitchell, 2021, p. 6)

⁶⁹ No original: *When putting cognitivist principles to the test in the design of intelligent applications, user interfaces, or efficient robots that must deal with the real world, it became apparent that more attention needed to be paid to the body in direct physical interaction with the environment. (Di Paolo, Buhrmann & Barandiaran, 2017, p. 15).*

⁷⁰ “An *effector* is any device that affects the environment. Robots control their effectors, which are also known as end effectors. Effectors include legs, wheels, arms, fingers, wings and fins. Controllers cause the effectors to produce desired effects on the environment” Fonte: <http://www.electronicsteacher.com/robotics/robotics-technology/effectors.php>

Tendo em perspectiva alguns desses problemas (mas não necessariamente identificando-os exatamente dessa maneira), em 1972 Hubert Dreyfus publicou *O que os Computadores Não Podem Fazer* (*What Computers Can't Do*, reeditado em 1992 sob o título *What Computers Still Can't Do*). A obra procura abordar a cognição humana a partir de uma crítica à inteligência artificial. Segundo Di Paolo et al. (2017, p. 15), em resumo o que Dreyfus fez foi “chamar a atenção para o insuficiente embasamento das premissas do cognitivismo na experiência vivida”. Enraizado na fenomenologia da percepção, pode-se dizer que Dreyfus apontou anomalias (para usar o termo de Kuhn) do cognitivismo, quando explicitou questões da vida prática que não poderiam ser compreendidas meramente a partir de premissas cognitivistas. O filósofo mostrou que, enquanto programadores da IA trabalhavam para que computadores pudessem resolver problemas cognitivos, se os problemas cognitivos não são todos possíveis de ser resolvidos com computadores (ao menos computadores como os conhecemos e conhecíamos), então isso configuraria, de saída, algo impossível de se concretizar.

Gostaria de dar prosseguimento ao raciocínio aqui exposto apresentando algumas das proposições de Dreyfus, cujas ideias considero compatíveis com muito do que hoje tem sido desenvolvido no âmbito da cognição corporificada e enativa.

3.3.1

O que os computadores *não* fazem diz muito sobre os humanos⁷¹

Dreyfus, cujo trabalho nasceu na efervescência do MIT das décadas de 1960 e 70 – a era da *GOF AI* – é um dos mais conhecidos e reconhecidos críticos da IA. Trabalhando ao lado de desenvolvedores de sistemas, investigou questões centrais, muitas vezes evitadas por conveniência entre aqueles que mantinham elevado o seu otimismo no sentido de reproduzir artificialmente a inteligência humana. Os aspectos que Dreyfus levantou são nevrálgicos porque suscitam questionamentos sobre como nós, seres humanos, adquirimos a capacidade de desempenhar as mais diversas atividades que cotidianamente protagonizamos. Por vezes, nossos atos parecem simples, em outros momentos extremamente complexos, mas é fato que demonstramos a todo momento um repertório variado de maneiras de lidar com as mais diferentes situações. Como conseguimos ir ampliando, ao longo da vida, esse

⁷¹ Conteúdo publicado no artigo Leporace (2022b), podendo haver trechos idênticos aqui.

acervo que nos permite atuar nas mais variadas situações, cenários e contextos físicos, sociais e culturais – não raro inéditos e surpreendentes para nós? Essa nossa habilidade – encapsulada por Dreyfus (2014) no conceito de *skillful coping* – deve-se ao fato de que temos um corpo e, a partir desse corpo, nós moldamos o nosso comportamento, sentimos, *aprendemos* por meio da experiência do nosso corpo no mundo⁷². Daí vem a nossa peculiar “habilidade” de lidar com o mundo.

A ênfase que Dreyfus atribui à experiência liga-se à postura adotada pela fenomenologia, vertente filosófica que se interessa pelo mundo tal como o experimentamos, pela maneira como ele se manifesta fenomenologicamente: como é que o mundo aparece para nós? Trata-se de uma pergunta tipicamente fenomenológica. Isso não deve ser considerado dado, e sim tomar o espaço da própria investigação. Esse aspecto chama-se redução fenomenológica (*phenomenological reduction*), como esclarece Thompson (2007, p. 18). Neste caso, redução não quer dizer a substituição de um modelo teórico em favor de outro, mas uma reorientação, uma busca de um foco específico. Quando adotamos essa postura fenomenológica, partimos do princípio de que as coisas vieram antes de nós, não estamos duvidando disso; nem achamos que aquilo que está em torno de nós é uma ilusão. No entanto, devemos procurar focar não naquilo que as coisas são, de uma maneira que Thompson (2007, p. 19) define como *naïve* (ingênua), e sim na maneira como as vivenciamos, como elas aparecem para nós, e como se relacionam com a nossa subjetividade. Por conta disso, a redução fenomenológica adota uma perspectiva da primeira pessoa. Deve-se suspender uma imersão no mundo caracteristicamente desatenta, trocando-a por uma capacidade de perceber

⁷² A abordagem de Dreyfus para a experiência corporificada é considerada pelos enativistas Di Paolo, Rohde e De Jaegher (2010) uma perspectiva em que a corporificação realmente é levada a sério. Uma vez que eles pontuam que o termo “embodiment” tem sido usado de uma maneira trivial, por vezes muito superficial em relação à proposta enativista para a corporificação, é importante notar como eles reconhecem a abordagem de Dreyfus desta maneira. *Many modern accounts of cognitive activity already take experience seriously. For instance, Dreyfus’s defense of nonrepresentational skill acquisition (2002) is based on paying careful attention to the experience of undergoing a process of task improvement. As we make the journey from beginners to experts through practice, not only is skillful performance improved, but experience is also transformed. This is to be expected if embodiment in the enactive sense is taken seriously. If experience and the body-in-interaction were to relate to each other as two mutually external systems, we would expect either an unchangeable or a fleeting relation between our bodies and our experience. Instead we find a lawful relation of bodily and experience transformations.* (Di Paolo, Rohde & De Jaegher, 2010, capítulo 2, p. 44)

essa imersão, um tipo de meta-consciência (*meta-awareness*), ou uma consciência da consciência⁷³.

Ao adotarmos a postura de direcionar a nossa atenção à maneira como a realidade se apresenta para nós, devemos focar efetivamente nos fenômenos, nas aparências e nos sentidos, nos significados que as coisas têm para nós, na maneira como aparecem para nós. Consequentemente, precisamos, para isso, investigar como esses fenômenos são constituídos – não no sentido de serem *fabricados*, como já esclarecido, mas sim de como são *trazidos à nossa consciência*. Esse é o sentido que o termo *constituição* tem para a fenomenologia⁷⁴. Nesse contexto, a concepção de fenômeno, portanto, também ganha novo significado. Zahavi (2019, p. 14) explica que fenomenologistas rejeitam uma *two-world doctrine*, a qual defende a existência separada de objeto e fenômeno – isto é, o mundo que se apresenta para nós e que pode ser entendido por nós *versus* o mundo que realmente existe. Isso não significa que fenomenologistas não separam o que realmente existe daquilo que são meras aparências, mas que, para eles, essa distinção não é entre diferentes realidades: é uma distinção entre *modos de manifestação*.

Ao defender a ideia de *skillful coping* – contrastando-a com a forma como sistemas de computação funcionam e, assim, explicando a inviabilidade de que computadores alcançassem essa habilidade humana – Dreyfus adota essa postura fenomenológica. Não há separação entre sujeito e mundo, e é isso que permite que se tenha esse *engajamento hábil*. A fenomenologia preconiza que a mente humana sempre está voltada para algum objeto do mundo; na verdade, ela só existe *porque* se volta para objetos do mundo. O foco da fenomenologia está, então, não na estrutura da mente, mas na díade mente-mundo, o que se torna mais claro a partir do conceito de intencionalidade. Fenomenologistas desejam explorar e esclarecer a relação entre a mente e o mundo, sendo o caráter intencional da mente o responsável por levá-la a transcender o *self*. A consciência não está contida num certo lugar e

⁷³ Husserl não propôs tal enfoque de uma maneira apartada da ciência, mas para fundar um novo marco filosófico para a investigação científica. Seu projeto teórico baseou-se em uma “reapreciação radical da experiência como fonte de significado e conhecimento” (Thompson, 2007, p. 19). Isso é importante, porque muitos acham que a fenomenologia nos afasta da realidade e é anticientífica. Mas, não se trata de uma rejeição da epistemologia, mas de uma epistemologia alternativa.

⁷⁴ Faço uma observação que considero pertinente, aqui: é necessário esclarecer e insistir em alguns vocábulos, termos e conceitos da fenomenologia com esse cuidado, quando se deseja compreender a relação de tal vertente filosófica com as teses contemporâneas para a cognição.

nem é um local especial, devendo, em vez disso, ser definida em termos dessa sua abertura.

Di Paolo et al. (2017) explicam que nós tomamos como dadas – sem questionar sua complexidade – tarefas que fazemos cotidianamente, como pedalar, consertar a pia ou andar de ônibus ou metrô na hora do *rush*. Acontece que, quando nos envolvemos em ações tão corriqueiras quanto essas, entram em jogo várias escalas simultâneas de tempo (referentes a cada pequena sub-ação que compõe a ação como um todo); a manipulação sensível de ferramentas e artefatos a partir do contexto; a necessidade de que coordenemos ações, variáveis, objetos, pessoas e de que lidemos, ao mesmo tempo, com normas, regras, objetivos que podem até ser conflitantes e gerar diversos microconflitos paralelos; a nossa sensibilidade aos riscos envolvidos em cada situação, de acordo com aquilo que o ambiente oferece; as capacidades e demandas do corpo. Mesmo com todos esses aspectos entrando em ação ao mesmo tempo, seres humanos geralmente lidam bem com essas situações. Di Paolo et al. (2017) esclarecem que nós tomamos a nossa vida sensório-motora trivialmente, quando, na verdade, estamos imersos nessas atividades como criaturas corporificadas, o que é um tanto complexo. Não à toa, as atividades aparentemente mais simples podem estar bastante além da capacidade de sistemas artificiais sofisticados e até de teorias cognitivas também avançadas, já que são atividades altamente sensíveis ao contexto e cheias de fatores que as modulam.

A era atual é a do machine learning: máquinas que preveem e calculam com uma capacidade de processamento incomparável à humana. Mas os computadores continuam não podendo realizar certas tarefas, enquanto em outras eles nos superam; geralmente, no processamento simultâneo de grandes quantidades de dados. Já atividades para além dos cálculos matemáticos podem surpreender pelo desnível entre a aparente simplicidade com que configuram no repertório cognitivo humano e o significativo desafio que impõem para serem formalizadas nos sistemas artificiais. Para compreender esse abismo, o que pode ajudar a entender por que não somos somente cérebros (ou pelo menos não somente a concepção de cérebro que tem predominado) – e ainda o porquê de a cognição não se limitar a processamento de dados – Dreyfus conduziu o olhar para o fato de que o nosso repertório amplo de habilidades para atuar em tantos cenários e contextos físicos, sociais e culturais, não raro inéditos e surpreendentes para nós, teria relação direta com o corpo e com as emoções. Na verdade, *emergiria do corpo e das emoções*. Como seres

corporificados, estamos inseridos no mundo desde sempre, sendo esse acoplamento com o ambiente um fator condicionante da nossa própria existência enquanto seres cognoscentes.

3.4

Estendendo a discussão: a vida que não cabe em um tutorial

Parece haver um gap entre o tanto que conseguimos *explicar* ou *compreender* sobre aprendizagem e o que realmente *acontece* quando aprendemos algo. Isto é, entre a epistemologia e a ontologia da aprendizagem. Um exemplo frequentemente presente em livros de filosofia e filosofia cognitiva, e que tem tudo a ver com o que desejo expor neste capítulo, é o da atividade de amarrar um cadarço. Como você se torna *bom* em amarrar o cadarço do tênis? Quais os passos que definiria para essa atividade? Como sabe se ficou bem amarrado ou se, assim que começar a andar ou correr, o laço irá se desfazer, o sapato ficar frouxo e você eventualmente cair? Trata-se de uma atividade que mal tem como ser descrita num passo-a-passo. Essa, afinal, é uma tarefa que envolve as mãos, a visão, o tato, a força, os pés, a pele – o corpo, enfim, especialmente quando se fica de pé e verifica se o cadarço está bem amarrado. É uma ação que possivelmente dá errado diversas vezes até que finalmente dê certo. Ou seja, envolve uma (ou várias) experiência(s), tentativas práticas – todas recrutando o corpo e seus atributos. Assemelha-se um pouco a outras como andar de bicicleta ou cozinhar (por que, às vezes, mesmo quando seguimos uma receita, o prato sai diferente do que esperávamos? Por que algumas pessoas têm um “tempero irresistível” e, ao mesmo tempo, tão difícil de ser transmitido pelas sucessivas gerações de uma família, ainda que sejam deixadas anotações em velhos cadernos?). Outros exemplos, bastante cotidianos, seriam limpar objetos, ou até mesmo saber como agir num vagão de metrô lotado.

O que essas atividades têm em comum? Todas elas se mostram bastante difíceis de ser formalizadas em regras, ou refletidas em um passo a passo. Elas até podem ter um “como fazer”, mas todo esse “tutorial” será limitado ao envolver simples palavras e instruções. Imagine um passo-a-passo para andar de bicicleta:

- Suba na bicicleta;
- Segure firme o guidão;

- Comece a pedalar o quanto antes para evitar que a bicicleta caia; ela cai se você para de pedalar.

Esse tutorial diz bem pouco, talvez nada, sobre a ação de pedalar. A parte mais importante, o sentir, o experimentar, fica de fora, pois não há como expressar isso sem que se sinta. Não à toa, quando um pai ou mãe ensina seu filho, segura na própria bicicleta, com a criança já montada sobre ela e, com um leve empurrão, ajuda o pequeno a pedalar; mas, quem pedala é a criança: a criança, a bicicleta, o chão e o ar, juntos, formam um sistema em que as partes colaboram para que se obtenha movimento e equilíbrio. É por isso que, se alguém lhe pedir para ensinar a andar de bicicleta, você dificilmente imaginará essa lição longe de uma bicicleta propriamente dita, e que esteja num espaço plano em que se possa pedalar. No caso do cadarço, é preciso que o aprendiz amarre e sinta o sapato no pé para perceber se está bem atado ou não.

Para aprender a pedalar, como uma alternativa, a lição pode começar por uma bicicleta com rodinhas, permitindo que a criança experimente por conta própria a sensação de se equilibrar – que envolve, nesse caso, os pneus contra o chão, o vento nos cabelos, as mãos no guidão, os pés nos pedais. Como disse Dreyfus (2005) em uma aula na Universidade da Califórnia, Berkeley, a nossa experiência indica que o papel das regras, quando se aprende algo, é como o papel das rodinhas para quem aprende a pedalar: ainda que precisemos das rodinhas/das regras no início, devemos deixá-las de lado se desejamos nos tornar ciclistas habilidosos. Porém, assumir que as regras que um dia seguimos conscientemente se tornaram inconscientes seria como assumir que, quando aprendemos a pedalar a bicicleta, as rodinhas teriam se tornado invisíveis, quando na verdade elas se tornaram desnecessárias e foram, então, desatarrachadas, para que o ciclista seguisse livre.

3.4.1

O contexto importa: o problema da relevância⁷⁵

No MIT, Dreyfus trabalhava lado a lado com cientistas da computação focados em reproduzir as habilidades humanas em computadores. Para ele, foi significativo observar a equipe então liderada por Marvin Minsky⁷⁶ mobilizada para

⁷⁵ Os dois primeiros parágrafos desta seção constam integralmente em Leporace (2002b), no prelo.

⁷⁶ Minsky foi um dos fundadores do MIT Computer Science & Artificial Intelligence Lab. Foi um matemático, cientista da computação e pioneiro na IA. Morreu em 2016, tendo publicado dez anos

resolver o problema do conhecimento do senso comum. Isso aconteceu nos anos de 1970, num momento que vinha marcado pelo otimismo na IA. Minsky declarava que teríamos máquinas com inteligência comparável à de humanos em 30 anos – em referência ao filme *2001, uma Odisseia no Espaço*, de Stanley Kubrick (Dreyfus, 2005, p. 48). Mas como seria possível ensinar aos computadores aquilo que parece que *todo mundo* sabe sem precisar que lhe seja ensinado? Como fazê-los entender histórias, referências ou argumentos que, quando ouvimos, parece que compreendemos de imediato? As sutilezas do mundo: aquilo que não costumamos colocar em palavras, que é difícil de explicar, mas que vamos absorvendo na medida em que experimentamos situações.

Quando se deparou com essa questão, Minsky (1974, 2007⁷⁷) acreditava que o obstáculo que veio a ficar conhecido como *commonsense knowledge problem* ou *problema do senso comum* seria resolvido a partir do armazenamento de milhões de regras e fatos nas máquinas, para que elas pudessem selecioná-los nas diferentes situações em que fossem solicitadas. Dreyfus, porém, entendeu que a solução não era essa. Parecia-lhe que a questão chave seria determinar quais elementos seriam relevantes em cada situação ou contexto⁷⁸. Essa seria a real capacidade de agir a partir do senso comum. Consequentemente, um computador não poderia chegar a *raciocinar* como um ser humano simplesmente depois de internalizar uma quantidade gigantesca de regras. E os impedimentos para isso estão, justamente, no fato de que o caminho que humanos seguem para a aprendizagem vai muito além da aquisição de regras; nós até podemos conhecer algumas delas, para que nos ajudem a começar a aprender algo; mas precisamos, sobretudo, saber para que as regras servem (e precisamos de tempo para aplicá-las e de sentir o que sentimos quando as aplicamos para de fato aprender). É preciso compreender, ainda, que essas regras podem mudar, quando o contexto muda. Humanos têm a capacidade de agir nessas mudanças; de selecionar estratégias para saber o que fazer para

antes seu último livro, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. Informações de seu obituário na página do MIT: <http://news.mit.edu/2016/marvin-minsky-obituary-0125>.

⁷⁷ Ainda que o diagnóstico de Minsky acerca disso tenha adquirido nova forma ao longo de sua trajetória como cientista da IA, em seu último livro ele ainda dedicava atenção ao problema do senso comum – definindo-o como “os tipos de fatos e conceitos que a maioria de nós sabe” (Minsky, 2007, p. 164) – e indicando que as máquinas não podem aprender pela experiência justamente porque lhes falta esse senso.

⁷⁸ Problema apontado também em Leporace et al. (2019, pp. 41-42), em capítulo sobre representações mentais na mente estendida e na mente enativa.

resolver os variados problemas que aparecem. Enquanto isso, no caso de um computador, uma mudança de contexto pode deixá-lo simplesmente sem resposta.

O *frame problem* é uma versão do já mencionado problema do senso comum: o problema da relevância, em que o computador fica sem saber como resolver quando trabalha com listas de situações desassociadas de necessidades, cenários, situações. O nome dessa vertente do problema advém de uma estrutura que Husserl havia chamado de “frame” em *Experience and Judgement*.

Minsky sugeriu como solução que os programadores de IA pudessem usar descrições de situações típicas como ir a uma festa de aniversário para listar e organizar aqueles, e somente aqueles fatos que normalmente seriam relevantes. Ele sugeriu uma estrutura de características essenciais e atribuições padrão - uma estrutura que Husserl já havia proposto e chamado de “frame”. (Husserl, 1973, p. 38) [...] (Dreyfus, 2007, p. 248)⁷⁹

Endereçar o problema criando um sistema com um vasto repertório de *frames*, ou quadros, como Minsky propôs no MIT, levaria a um outro problema: o regresso infinito. Afinal, para cada quadro correspondente a um fato deveria haver outro, correspondente a uma situação, e assim por diante, sem que se vislumbrasse um fim. Para Dreyfus, mais que um problema, isso indicava uma abordagem equivocada:

Um sistema de frames não configura uma situação; portanto, para identificar os fatos possivelmente relevantes numa situação corrente, seria necessário um frame para reconhecer essa situação, etc. Assim, pareceu-me óbvio que qualquer programa de IA que usasse frames seria pego em um retrocesso infinito de frames para o reconhecimento de fatos relevantes, e que, portanto, o problema do armazenamento e da recuperação do conhecimento de senso comum não era apenas um problema; era um sinal de que algo estava seriamente errado com toda a abordagem. (Dreyfus, 2007, p. 248)⁸⁰

A *Stanford Encyclopedia of Philosophy* (2004)⁸¹ fornece um exemplo ilustrativo para se compreender as dificuldades envolvidas no problema do

⁷⁹ No original: *Minsky suggested as a solution that AI programmers could use descriptions of typical situations like going to a birthday party to list and organize those, and only those, facts that were normally relevant. He suggested a structure of essential features and default assignments—a structure Husserl had already proposed and called a “frame” (Husserl, 1973, p. 38) [...] (Dreyfus, 2007, p. 248)*

⁸⁰ No original: *[A] system of frames isn’t in a situation, so in order to identify the possibly relevant facts in the current situation one would need a frame for recognizing that situation, etc. It thus seemed to me obvious that any AI program using frames was going to be caught in a regress of frames for recognizing relevant frames for recognizing relevant facts, and that, therefore, the commonsense knowledge storage and retrieval problem wasn’t just a problem; it was a sign that something was seriously wrong with the whole approach. (Dreyfus, 2007, p. 248)*

⁸¹ Ver <https://plato.stanford.edu/entries/frame-problem/>

conhecimento do senso comum e a tentativa de se endereçar essa questão por meio de *frames*. Ainda que o exemplo utilize um robô corporificado em vez de uma máquina no formato de um computador tal como conhecemos, parece válido trazê-lo para ampliar a compreensão. O exemplo nos pede para imaginar um robô que precisa desempenhar, todos os dias, a mesma tarefa: fazer uma xícara de chá. Imagine que a xícara tem que ser retirada do armário pelo robô, e que a localização da xícara está representada na base de dados do robô junto a outros fatos que representam inúmeras outras características de uma situação em andamento, como a temperatura ambiente, a configuração dos braços do robô, a cor do pote de chá. Depois de pegar a xícara e a retirar do armário, a base de dados do robô precisa ser atualizada. Afinal, a localização da xícara mudou. Quando isso acontece, quais outras frases do código daquele robô (sua programação) precisam ser modificadas? Se a temperatura do ambiente não mudou, e nem a localização do pote de chá, mas uma colher estava dentro da xícara, a localização dessa colher mudou junto com a mudança de lugar da xícara, e, sendo assim, também precisa ser atualizada, por exemplo. Como o robô poderia limitar o escopo das proposições que deveria reconsiderar à luz dessas ações? Num robô mais simples, isso não parece tão complicado, porque sua base de dados poderia ser examinada, dado por dado, até que ele identificasse o que é que precisa mudar. Porém, se o robô tivesse uma base de dados extensa e complexa, a quantidade vasta de fatores a serem analisados e modificados a cada vez que algo muda na situação “vívda” pelo robô se tornaria algo moroso e extenso.

Se nós, humanos, operássemos assim, na base de *frames*, avaliando cada item do ambiente em separado, absorvendo cada elemento de maneira distinta do outro, para depois entender a relação entre eles, provavelmente iríamos “travar” diante de certas situações, especialmente quando tentássemos empreender a atitude adequada e não obtivéssemos sucesso. Mesmo que as listas e quadros disponíveis se associassem a alguns contextos, esses contextos mudariam, já que dependem dos eventos a nossa volta a todo o tempo; então, não saberíamos o que fazer, da mesma forma. Por conta disso, o *frame problem* parece apropriado para pensarmos em como a totalidade de uma situação, e a nossa capacidade de *perceber* essa totalidade, é que nos ajuda a saber como agir. Isso vale para uma melodia – as notas adquirem valor na medida em que constituem essa melodia e são percebidas como partes da música; vale para a percepção de objetos – se eu sei que estou olhando

para uma maçã de cera e ela tem uma camada nebulosa sobre sua casca, posso associá-la a poeira, enquanto se fosse uma maçã fresca eu poderia pensar que a camada era de umidade. Dreyfus menciona, ainda, como pode ser difícil para uma máquina compreender, por exemplo, o tom empregado por um sujeito ao falar determinada frase. Esse é um aspecto que pode mudar o sentido todo daquilo que se diz – e que para nós não seria tão difícil captar, em geral (Dreyfus, 1992, p. 238).

Contando a anedota de uma mulher que recebe a ligação telefônica de um homem, Minsky (2006, p. 165) fornece um bom exemplo de senso comum. Ele relata que, ao ouvir o telefone tocar, uma mulher, que está em casa, vai atendê-lo; o homem, do outro lado da linha, explica a ela sobre um processo químico,

respondendo a uma pergunta que ela lhe havia feito; ele a orienta a ler um determinado livro, e diz que irá levar esse livro a ela em breve, pois terá que ir a um local próximo da casa dela; a mulher agradece ao homem e termina a ligação. O homem, depois de um tempo, aparece e lhe entrega o livro. Depois de contar essa curta história, Minsky lista uma série de entendimentos subjacentes aos envolvidos; por exemplo: quando a mulher ouve o telefone tocar, ela reconhece que o toque significa que alguém quer falar com ela; ela atravessa a sala para pegar o telefone e o coloca na orelha; ambos sabem usar um telefone e sabem que estão em locais diferentes; ela não ficará surpresa quando ele chegar para lhe entregar o livro porque ele diz que em breve estará pela sua vizinhança, e assim por diante. Percebe-se, pela história ilustrada por Minsky, que a mulher tem uma apreensão completa da situação, assim como o homem. Isso vale para a percepção física dos detalhes envolvidos, também: ao pegar o telefone, a mulher não vê o aparelho todo, mas uma parte dele. Mesmo assim, percebe o telefone como um todo. E, mesmo antes de alcançar o aparelho para atendê-lo, ela antecipa a forma como o objeto se moldará à sua pegada (Minsky, 2006, p. 166).

Exemplifico, agora, essa apreensão do contexto a partir de um cenário típico de sala de aula. Em uma comparação entre *e-learning* e ensino presencial, Dreyfus (2009, p. 60) apresenta o relato de um professor sobre como presencialmente ele consegue observar seus alunos de maneira mais orientada – se vários alunos balançam a cabeça concordando com a pergunta feita por um deles, por exemplo, ele pode perceber que outros provavelmente têm a mesma dúvida e possivelmente encaminhar esclarecimentos a partir dessa percepção; pode perceber se os alunos estão dispersos, sonolentos, estabelecendo conversas paralelas etc.; já com uma

aula transmitida por câmera, estando cada um em seu ambiente particular, essa apreensão da situação fica prejudicada.

Podemos pensar em outros exemplos para além desses. Quando temos familiaridade com um certo conjunto de objetos, que geralmente para nós aparecem juntos, não nos surpreendemos de vê-los naquele contexto; isso acontece com o vidro de azeite na hora de uma refeição – mesmo que não tenha rótulo, se estiver sobre a mesa em uma garrafa de vidro com um bico fino e tiver uma determinada coloração amarelo-esverdeada vamos perceber azeite ali, e não detergente, por exemplo (claro, podemos errar; a visão pode nos enganar, como Descartes diria...); se estamos numa aula de pintura e nos deparamos com manchas pretas em nossa roupa, percebemos ali tinta, e não molho *shoyo* (ainda que o molho possa ser usado como tinta, se o artista desejar inovar!). Se estivéssemos caminhando e víssemos uma janela sobre o chão da rua, isso talvez nos deixasse confusos por uns instantes, mas dificilmente ficaríamos confusos simplesmente ao ver janelas se estivéssemos dentro de uma casa ou olhando para ela, pois isso é esperado, faz parte do cenário usual; se estamos navegando em um site e vemos uma seta para o lado, percebemos que clicar nela pode significar passar de página, e assim por diante. De um modo geral, quando não sabemos alguma coisa, fazemos associações e podemos resgatar o contexto original de tais objetos, passando então a compreendê-los nesse novo contexto.

Nos exemplos de Dreyfus, ele procura se limitar à percepção, por ser uma atividade envolvida no nosso conjunto cognitivo mais básico e, ao mesmo tempo, aparentemente mais instigante – e não necessariamente mais simples. Ressalto que, na perspectiva de Dreyfus, o *frame problem* não existe para seres humanos, porque funcionamos de outra maneira: nosso acoplamento com o mundo nos garante a capacidade de contextualizar a todo instante, no desenrolar das atividades que vamos desenvolvendo. É intrigante pensar que logo uma capacidade aparentemente tão trivial do ser humano é aquela que lança um dos maiores desafios à formalização de tarefas nas máquinas, contribuindo para o *gap* entre o funcionamento humano e o artificial.

Por que o problema do contexto é tão relevante quando se trata de buscar um paradigma para a cognição que vai além do cérebro? Basicamente porque a habilidade de estar no mundo e agir de acordo com cada situação recruta muito mais do que processamento de informações pelo cérebro, com consequente geração de

outputs. Mesmo saindo do pressuposto da *GOFAI*, em que predominavam as extensas e detalhadas regras de programação, e entrando no dos atuais sistemas de machine learning, baseados em algoritmos, muito do que é essencialmente humano é ainda deixado de fora, por mais capazes que sejam esses recursos. Vamos olhar para um exemplo que tem a ver com a atividade de cozinhar: o de uma receita. Qual seria a diferença entre um algoritmo e a receita de um prato?

Um algoritmo não é qualquer conjunto de instruções: eles devem ser precisos e inequívocos o suficiente para serem executados por um computador. Por exemplo, uma receita culinária não é um algoritmo porque não especifica exatamente em que ordem fazer as coisas ou exatamente o que cada etapa é. Exatamente a quanto açúcar corresponde uma colher? Como todos que já experimentaram uma nova receita sabem, segui-la pode resultar em algo delicioso ou uma bagunça. Em contraste, um algoritmo sempre produz o mesmo resultado. Mesmo que uma receita especifique precisamente meia onça de açúcar, ainda não estamos perdidos porque o computador não sabe o que é açúcar, ou o que é uma onça. Se quiséssemos programar um robô de cozinha para fazer um bolo, teríamos de dizer a ele como reconhecer o açúcar assistindo a vídeos, como pegar uma colher e assim por diante. (Ainda estamos trabalhando nisso.) O computador precisa saber como executar o algoritmo até ligar e desligar transistores específicos. Portanto, uma receita culinária está muito longe de um algoritmo. (Domingos, 2015, p. 3)⁸²

Pedro Domingos é pesquisador em ciência da computação. Como ele coloca, o que falta à receita culinária em relação ao algoritmo é precisão; para que a máquina possa seguir esses passos à risca, cada detalhe precisa ser fornecido. Por outro lado, é interessante pensar: a atividade humana não é 100% precisa. Aparentemente, a falta de precisão é que pode preencher certas lacunas impreenchíveis pelos sistemas de IA. Usamos uma “pitada” de sal, “um pouco” de queijo ralado, e, de forma quase espontânea, por vezes sabemos as quantidades básicas de uma receita, sem que fiquemos paralisados sem conseguir progredir diante da possível ausência de alguma informação. Um sistema artificial, diante dessa falta, fica simplesmente sem saber o que fazer. O que uma receita feita por

⁸² No original: *An algorithm is not just any set of instructions: they have to be precise and unambiguous enough to be executed by a computer. For example, a cooking recipe is not an algorithm because it doesn't exactly specify what order to do things in or exactly what each step is. Exactly how much sugar is a spoonful? As everyone who's ever tried a new recipe knows, following it may result in something delicious or a mess. In contrast, an algorithm always produces the same result. Even if a recipe specifies precisely half an ounce of sugar, we're still not out of the woods because the computer doesn't know what sugar is, or an ounce. If we wanted to program a kitchen robot to make a cake, we would have to tell it how to recognize sugar from video, how to pick up a spoon, and so on. (We're still working on that.) The computer has to know how to execute the algorithm all the way down to turning specific transistors on and off. So a cooking recipe is very far from an algorithm. (Domingos, 2015, p. 3)*

um ser humano, até mesmo na base do improviso, tem em comum com a atividade de amarrar o sapato, que mencionei? Todas se mostram bastante difíceis de ser formalizadas em regras, ou refletidas em um passo a passo. Elas até podem ter um “como fazer”, mas todo esse “tutorial” será limitado ao envolver simples palavras e instruções. Somente sentindo, experimentando, o sujeito que aprende a fazer aquela atividade será capaz de efetivamente dizer que aprendeu a realizá-la.

Ainda que esta não seja uma investigação que se proponha exatamente a fazer uma comparação entre humanos e sistemas artificiais, esse é um aspecto que acaba por tangenciar o tema e inevitavelmente surge como um ponto a ser observado. Especificamente, para fins de compreender como a aprendizagem humana pode ser diferente da aprendizagem de máquina e o que isso pode significar, importam as atividades que possivelmente sejam ou não sejam formalizáveis de maneira sintética. Esse é, afinal, um exercício que nos leva a pensar nas características possivelmente inerentes à atividade cognitiva humana, já que por vezes o que parece mais difícil de se tornar uma regra formal entra na categoria das coisas que classificamos como mais descomplicadas entre as nossas atividades.

3.5 Conclusão do capítulo

Procurei demonstrar, neste capítulo, que o paradigma predominante para a mente humana ainda é o da mente computacional, e de que maneira seus pressupostos se estendem às tecnologias de IA, inclusive aquelas baseadas em machine learning. O ponto central aqui é o de que tecnologias concebidas a partir da premissa de que cérebro é sinônimo de mente, ainda que admitam a importância do corpo, são permeáveis a práticas educacionais que, igualmente, apesar de incluir o corpo nos processos de ensino e aprendizagem, ainda o concebem como um elemento periférico. Quando se parte da premissa de que é o cérebro que aprende, tais tecnologias mostram-se permeáveis a esse contexto porque vêm, afinal, ao encontro dessa concepção de aprendizagem cérebro-centrada que permanece em destaque, se dispondo a endereçar os nós educacionais que surgem nesse contexto. Um dos problemas de restringir a mente ao cérebro, e a inteligência a símbolos sintetizáveis, paraleliza com o problema de tomar a aprendizagem como algo que poderia estar circunscrito a algoritmos. Deve-se buscar compreender que se, por um

lado, a mente não se limita ao cérebro, por outro a aprendizagem tende a não se limitar a aquilo de que os algoritmos podem dar conta.

Tendo o nascimento da inteligência artificial como área de pesquisa ocorrido na mesma época em que a ciência cognitiva se constituiu como campo, não há como ignorar a maneira intercambiada como um informa o outro. Mas, apesar de o nascimento da IA datar de época semelhante ao despontar da ciência cognitiva, esta última conta com um movimento que hoje testemunha a emergência de um novo paradigma. Enquanto isso, a IA ainda se mostra incipiente no que diz respeito à absorção das premissas mais fortes da cognição corporificada. Sendo assim, se a ciência cognitiva alimenta a educação – já que pesquisas sobre a cognição geram, ou devem gerar, impactos e avanços nos conceitos e ideias envolvidos na psicologia da educação, na filosofia, na antropologia, na sociologia, na compreensão da estética e da ética na educação, por exemplo, e ainda nas maneiras de se compreender e conceber as tecnologias educacionais – torna-se importante olhar nessa direção quando se deseja uma perspectiva crítica para as tecnologias digitais com fins educacionais.

Afinal, se não houver um olhar em direção a um novo possível paradigma, continuaremos a ver essas tecnologias sendo aplicadas ao contexto educacional (formal ou não) a partir de premissas que talvez devessem ser revistas. Isto é, corre-se o risco de promover modificações casuais, superficiais, sem se alterar algo nas bases, nos fundamentos da educação, algo que poderia ser prioritário. Além disso, numa era em que os sistemas de machine learning por vezes se mostram capazes de superar a capacidade humana, particularmente quando se trata de fazer cálculos rápidos e simultâneos, vale procurar olhar para o que algoritmos não podem fazer. O que esses sistemas não conseguem realizar ou deixam de fora talvez seja aquilo de que a educação menos pode prescindir. Essa é a frequência da complementaridade entre as tecnologias e o humano, que é bastante diferente da direção que se toma quando se preza pela substituição humana por recursos digitais.

Com seus avanços e também em seus “invernos” – as fases em que as pesquisas não correm tão bem quanto esperado – a IA acaba revelando muito sobre como funcionamos e também sobre como não funcionamos. As pesquisas tanto em IA como em robótica vêm ajudando a enfatizar a ideia de que a mente não está restrita ao cérebro e, portanto, examiná-lo como um elemento isolado pode não ser o caminho para se compreender os processos mentais de maneira ampla. Isso é até

um pouco irônico, no caso da IA em forma de machine learning em particular; pois essa é uma vertente que foca essencialmente numa arquitetura inspirada no cérebro e se volta para desenvolver cérebros, quando aplicada a plataformas educacionais, mas faz tempo que a própria IA ajudou a revelar empiricamente que não somos somente os nossos cérebros. Por que as tecnologias digitais baseadas em machine learning deveriam ter como único foco o desenvolvimento desse órgão, circunscrevendo assim os processos de aprendizagem, se a ciência cognitiva e os seus avanços nos indicam caminhos para conceber a aprendizagem de um modo mais amplo?

Ao propor que a aprendizagem seja tomada mais amplamente, me refiro a alguns fatores; por exemplo: que o desenvolvimento da pesquisa em cognição vem mostrando que não funcionamos somente na base de computação, ou ao menos não funcionamos somente a partir de computação ocorrida no cérebro; que a semântica ou o significado das coisas para os seres cognoscentes precisa de mais do que manipulações sintáticas para emergir; que talvez não necessitemos de representações mentais, ou ao menos possamos prescindir delas em algumas atividades cognitivas; que não somos somente cérebros, pois a atividade cognitiva precisa do corpo e de sua atividade no mundo para se desenrolar; que percepção e ação acontecem não uma após a outra, mas simultaneamente e em muitas escalas de tempo concomitantes.

Esses são aspectos os quais vêm sendo tratados pelas vertentes que, no decorrer dos anos, desafiam o cognitivismo primário, tradicional. Os temas a que me referi interessam, todos, a esta pesquisa que aqui apresento. A descrição da mente como algo computacional, uma receptora de *inputs* a serem transformados em *outputs* por meio do processamento de informações, pode ser tomada como a de um sistema cognitivo estático, previsível, que se utiliza de algo preexistente para gerar conhecimento. Aquilo que está no mundo se torna alimento para a cognição, sendo digerido pelo nosso “hardware”. Alimento pode ser tomado como algo pronto, que se sabe previamente o que é; da mesma maneira, consumir *input* para gerar *output* é um processo que requer linearidade, numa previsível sequência em que um vem depois do outro. Por mais que haja uma variação do cognitivismo na forma do connexionismo, ainda estamos falando essencialmente do cérebro e daquilo que acontece dentro dele quando absorve *inputs* do mundo para devolvê-los na forma de *outputs*. Admitimos que o cérebro precisa do restante do corpo para

funcionar, mas o peso da atividade cognitiva permanece concentrado na atividade cerebral, como se todo o restante fosse um apoio. A importância do corpo no mundo para a cognição será tratada em maior profundidade quando forem apresentadas as bases essenciais do enativismo.

Ainda que a ciência cognitiva tenha caminhado para reunir, como área de investigação, outras propostas para além do original programa cognitivista, boa parte da chamada teoria computacional da mente persiste como o paradigma predominante para a cognição na atualidade. No entanto, isso não significa que não haja mudanças significativas em curso. Pouco antes de, em 1992, Dreyfus reeditar (com muitos acréscimos importantes sobre redes neurais, por exemplo) seu *What Computers Can't Do*, desta vez sob o título de *What Computers Still Can't Do*, foi lançado em 1991 (por Francisco Varela, Evan Thompson e Eleanor Rosch) *The Embodied Mind*. A obra é uma das grandes responsáveis por plantar as sementes da cognição corporificada e enativista. Rolla (2021) comenta sobre o que a obra significa para a história da ciência cognitiva:

O enativismo foi apresentado originalmente no grandioso livro *The Embodied Mind* (1991/2016) escrito por Francisco Varela, Evan Thompson e Eleanor Rosch como alternativa ao cognitivismo de velha guarda. Embora essa obra seja sob certos aspectos muito programática, não é por acaso que ela tenha se tornado o *locus classicus* da filosofia das ciências cognitivas que já vinha se desenhando paralelamente ao declínio do paradigma cognitivista. Nela, os autores virtuosamente reúnem influências de inteligência artificial, fenomenologia, pragmatismo americano, teorias de sistemas dinâmicos e teoria dos sistemas desenvolvimentais que culminaram, como veremos, em uma concepção da mente como *enativa, corporificada e situada*. Isso marca uma ruptura radical com o antigo cognitivismo, que tratava as relações entre organismo e ambiente, bem como a dimensão experiencial da cognição, como meros subprodutos da computação interna sobre representações. Com efeito, não é exagero classificar *The Embodied Mind* (doravante, TEM) como um *novo paradigma*, dada a importância renovada que a ação e a corporeidade receberam a partir desse livro, o que o caracteriza como um dos principais responsáveis pela *virada pragmática* nas ciências cognitivas (Engel et al., 2013). (Rolla, 2021, p. 47-48, grifos do autor).

No próximo capítulo, apresento a proposta do enativismo para a cognição, introduzindo suas ideias e conceitos principais. Esses estudos podem trazer a bordo um novo paradigma para a cognição. A proposta é que sejam identificados no enativismo preceitos que possam contribuir para que se enxergue também um novo possível paradigma para a aprendizagem humana⁸³.

⁸³ Ver Bannell, Leporace e Santos, 2021.

Only intertwining these two perspectives, the biological and the phenomenological, can we gain a fuller understanding of the immanent purposiveness of the organism and the deep continuity of life and mind.
(Thompson, 2007, p.148).

Se, em 1970, Marvin Minsky presumia que existiriam máquinas comparáveis a humanos em três décadas, em 1990 o roboticista Hans Moravec – ainda hoje trabalhando na Universidade de Carnegie Mellon, nos EUA – acreditava que robôs “com inteligência humana” seriam comuns em 50 anos (Moravec, 1990, p. 6). Mais cauteloso talvez tenha sido Rodney Brooks – roboticista australiano também ainda em atividade, dirigindo o Laboratório de Inteligência Artificial e Ciência da Computação do MIT – quando, em 1987, escreveu que “ninguém mais fala sobre replicar a gama completa da inteligência humana”⁸⁴. Ele, entre outros, acreditava que

(...) a inteligência em nível humano é complexa demais e pouco compreendida para ser corretamente decomposta em subpartes no momento e (...) mesmo que conhecêssemos essas subpeças, ainda não saberíamos quais as interfaces corretas entre elas” (Brooks, 1987, p. 139)⁸⁵.

Já se passaram mais de trinta anos desde o palpite de Moravec. Em mais duas décadas, teríamos de fato os robôs imaginados por ele? No que concerne à persistência dos desafios impostos pela natureza da cognição humana à IA, parece que não. Por mais que tenhamos máquinas capazes de nos superar em atividades que envolvem cálculos, com o processamento simultâneo de um vasto número de informações, ou robôs capazes de carregar pesos que humanos não poderiam, em fábricas de automóveis por exemplo, estamos distantes de uma IA generalista. O problema do senso comum, o qual apresentei no capítulo anterior, não apenas é um obstáculo para a IA – contribuindo para aumentar o *gap* entre humanos e sistemas artificiais – como está longe de ser o único entrave ainda sem solução, no que diz respeito a sistemas artificiais, a remeter a esse *gap*.

⁸⁴ No original: *No one talks about replicating the full gamut of human intelligence anymore.* (Brooks, 1987, p. 139)

⁸⁵ No original: *[H]uman level intelligence is too complex and little understood to be correctly decomposed into the right subpieces at the moment and (...) even if we knew the subpieces we still wouldn't know the right interfaces between them.* (Brooks, 1987, p. 139)

Enquanto uma gama de tarefas consideradas difíceis para a cognição e a inteligência humana são desempenhadas com sucesso por sistemas artificiais – os quais se mostram capazes de jogar xadrez, se saem bem nos chamados testes de inteligência, resolvem tarefas de álgebra, teoremas e daí por diante – outras atividades, que acabamos por conceber como triviais ou ordinárias, são ainda consideradas muito complexas para serem executadas por sistemas artificiais. Encontram-se no grupo dessas tarefas algumas que mesmo um bebê pode protagonizar; dentre elas, ações relacionadas à percepção e à mobilidade. Curiosamente, essas questões desafiadoras para a IA foram aglutinadas sob o rótulo do chamado “Paradoxo de Moravec” (Dellermann et al., 2019) – pelo próprio Moravec, por Marvin Minsky e por Brooks. O trecho a seguir resume a contradição expressa pelo paradoxo:

Infelizmente para os robôs humanoides, os computadores encontram seu pior resultado quando tentam fazer as coisas mais naturais para os humanos, tais como ver, ouvir, manipular objetos, aprender línguas e raciocinar no senso comum. Essa dicotomia – máquinas a fazer bem as coisas que os humanos acham difíceis, ao mesmo tempo que fazem mal o que é fácil para nós – é uma pista gigantesca para o problema de como construir uma máquina inteligente. (Moravec, 1988, p. 9)⁸⁶

O trabalho de investigação da cognição humana que o enativismo procura fazer pode indicar novas diretrizes para desafios que, como esses, há muitos anos foram expostos pela IA ao encará-los. É justamente a partir da questão do senso comum que os autores propõem as bases do enativismo: fazendo menção ao jogo de xadrez, no qual as máquinas de IA se saem bem, eles explicam que uma abordagem à cognição focada na resolução de problemas funciona para situações em que é possível e até relativamente fácil prever todos os possíveis estados – no caso do xadrez, posições no tabuleiro e jogadas. No entanto, não é uma abordagem tão produtiva quando se trata de domínios ou tarefas não tão bem definidos. Para isso, dão o exemplo de um robô que precisa dirigir em uma cidade movimentada. Nesse caso, por mais que o espaço possa ser minimamente definido a partir de certos aspectos, como outros carros, sinais de trânsito etc., não se trata de um ambiente “fechado”, com fronteiras bem delimitadas. Trata-se de um cenário que,

⁸⁶ No original: *Unfortunately for humanlike robots, computers are at their worst trying to do the things most natural to humans, such as seeing, hearing, manipulating objects, learning languages, and commonsense reasoning. This dichotomy—machines doing well things humans find hard, while doing poorly what is easy for us—is a giant clue to the problem of how to construct an intelligent machine. (Moravec, 1988, p. 9)*

por ser precariamente demarcado, não tem um começo ou um fim exato. Os movimentos são contínuos nesse espaço, cuja estrutura forma-se em vários níveis aglutinados (exemplos em Varela, Thompson & Rosch, 2016, p. 147). Citando Dreyfus, os autores explicam que “esse conhecimento do senso comum é difícil, talvez impossível de ser traduzido como conhecimento proposicional, explícito – ‘knowledge *that*’ versus ‘knowledge *how*’ – jargões da filosofia que indicam algo como *saber o que fazer e saber como fazer*” (Varela, Thompson & Rosch, 2016, p. 148).

Concomitantemente, assim, a cognição enativa abre perspectivas alternativas a aquelas mais tradicionais para a compreensão da cognição humana. Por meio de um conceito central, o de *sense-making* (que será apresentado neste capítulo, junto a outros conceitos de base da abordagem), enativistas encapsulam a ideia da cognição no seu estado mais bruto, por assim dizer, definindo-a como a capacidade do agente de fazer o entorno aparecer para ele. As coisas que existem no meio em que o agente vive só *fazem sentido* para ele quando, ao agir no mundo, ele percebe o ambiente, se deixa *afetar* pelos elementos ao seu redor. Assim, esses elementos

passam a *significar algo* ou a *ter algum sentido* para esse agente – isto é, efetivamente *aparecem* ou *surgem* para ele. Isso só é possível por conta do corpo, a interface por meio da qual o organismo em contato com mundo, afeta o mundo e se deixa afetar por ele. É nesses termos que se pode dizer que o mundo emerge para o agente a partir de sua capacidade corporal de se movimentar no ambiente; e é precisamente essa habilidade que começa a caracterizar um ser como cognoscente.

Estamos falando de capacidades perceptivas e de habilidades sensório-motoras que começamos a desenvolver ainda bebês; ações como essas instigam os desenvolvedores de sistemas artificiais, por seguirem sendo complicadas mesmo para sofisticados robôs. Nesse leque de dificuldades, cuja chave é o corpo no mundo – no enativismo, a ideia da corporificação ganha novo e substancial contorno – um outro aspecto a ser observado é o da experiência, que ganha aqui um novo sentido, sendo permeada pela capacidade de se deixar afetar pelos elementos do mundo. Para dar um exemplo didático sobre como se pode ou não ter experiência, sistemas de machine learning podem classificar inúmeros chocolates⁸⁷ – como pretos,

⁸⁷ Este exemplo foi dado no dia 5 de novembro de 2021 em reunião online do *Reading Group Filosofia da Inteligência Artificial e da Computação* da Universidade de Lisboa (de que participei por alguns meses; reúne pesquisadores de várias universidades, principalmente de Portugal) pelo

brancos, doces, amargos, com castanhas, com uma determinada porcentagem de cacau ou de açúcar; podem “dizer” se contêm leite integral ou desnatado, adoçantes e daí por diante, mas esses sistemas não “sabem” realmente qual é o sabor, o gosto, ou a textura que um chocolate tem. Nunca efetivamente *experimentaram* chocolate algum⁸⁸. Também não sentem nada quando classificam as pessoas de determinadas maneiras, porque não as conhecem, de fato. Não são capazes de ficar felizes ou chateados quando atingem ou não determinado objetivo, quando se saem bem ou mal numa tarefa e quando provocam alguma reação nas pessoas. Computadores processam e executam. Quando se trata de sentir o que sentimos, podem apenas simular.

Além da questão da *corporificação* – a qual, no enativismo, é colocada de uma maneira específica (os enativistas fazem questão de ressaltar que não basta incluir o corpo na cognição, mas que a cognição deve ser concebida antes de tudo *a partir do corpo*; também desenvolveram pilares a partir dos quais definem o que seriam corpos, os quais apresentarei adiante); do conceito de *sense-making*, proposto por autores de uma determinada vertente do enativismo (como irei explicar adiante) e do conceito de *experiência* tal como é concebido nesse contexto, são indispensáveis à compreensão da proposta enativista aqui introduzida os conceitos de *autonomia* e *emergência*, dos quais também tratarei. Ao apresentá-los, neste capítulo, procuro seguir a investigação sobre as razões por que o enativismo pode ser considerado um paradigma emergente para a ciência cognitiva e o que isso pode significar para a compreensão da aprendizagem humana em contraste com a aprendizagem de máquina.

prof. Vítor Santos, da NOVA Information Management School (NOVA IMS) da Universidade Nova de Lisboa, onde leciona disciplinas das áreas dos Sistemas de Informação e da Inteligência Artificial. Na ocasião, outro membro do grupo, prof. João Luís Cordovil, Coordenador Científico do CFCUL – Centro de Filosofia das Ciências da Universidade de Lisboa, comparou a analogia apresentada por Santos com a conhecida experiência mental da cientista Mary que, em um laboratório, consegue ser conhecedora de tudo, mas quando conhece a cor vermelho percebe que não haveria como saber o que é vermelho sem efetivamente ter a experiência perceptiva, sensorial, fenomenológica daquela cor. “Podemos dar à máquina (sem mente) os efeitos trazidos pelo chocolate, mas não a experiência do chocolate em si. Sem mente, elas mimetizam, copiam, utilizam artifícios sintéticos para ter algo o mais parecido possível com a experiência”, disse Cordovil, na ocasião.

⁸⁸ Ainda que existam sistemas capazes de emular sensações; de operar como se experimentassem.

4.1

(N)o caminho do imprevisível: percepção e(é) ação

Abordagens contemporâneas à mente e à cognição indicam que, como sistemas complexos e em constante processo de transformação, seja na atividade individual, seja nas constantes trocas com os outros, somos mais do que os nossos cérebros. Nomeadamente, somos mais do que cérebros individuais, meramente a processar informações oriundas do meio externo. Essas abordagens ajudam a explicar por que previsões construídas sobre o nosso passado vivido são – e tendem a ser sempre – limitadas em alguma medida. Curiosamente, a própria história das máquinas de ensinar dá pistas apenas parciais sobre o futuro desses dispositivos: como se poderia prever onde estaríamos hoje, com machine learning somado a big data? Quando pessoas, leigas ou não, num passado não tão distante levantaram questões acerca dos usos das máquinas de ensinar e da filosofia por trás delas, por exemplo, elas não apenas levantaram dúvidas sobre um certo aspecto como se fosse algo apartado de suas vidas: aquilo de fato lhes afetava, dizia respeito à educação de seus filhos ou delas mesmas, ou às instituições em que trabalhavam, enfim.

Sendo assim, ao apontar possíveis consequências decorrentes do uso desses dispositivos, as quais vislumbravam com sua intuição e com o repertório de experiências com que então contavam, esses críticos contribuíram para interferir no curso das coisas. Tanto que houve um “inverno” na história desses recursos, advindo de muitos fatores, mas também das críticas levantadas. Como se poderia adivinhar o futuro, se ele afinal estava ainda para ser construído pelas mesmas pessoas que viviam aquele presente e outras que ainda viriam?

Penso tais questões a partir de uma tese central em Maturana e Varela (2019)⁸⁹ que é característica do enativismo tal como tem sido proposto; trata-se da noção da inseparabilidade entre os seres humanos e o mundo, de forma que um possibilita e faz emergir o outro; não *somos* sem o mundo, e ele não *é* sem a nossa experiência e o nosso constante trabalho de atribuição de sentido e significado a aquilo que se encontra em torno de nós. “Todo ato de conhecer faz surgir um mundo” (Maturana e Varela, 2019, p. 31-32). O contexto, o entorno, o ambiente ou o meio, então, se constitui a partir dessas relações, no sentido de que emerge, isto é, vem à tona de fato a partir delas. Sendo assim, pode-se dizer que, em vez de haver

⁸⁹ Defendida na obra “A Árvore do Conhecimento”, primeira edição de 1987.

algo como um caminho previamente traçado à espera de sofrer interferências, há um caminho que se faz ao caminhar: eis um posicionamento tipicamente enativista. Sendo assim, quando a sociedade questiona algo, pode acabar por mudar o rumo das coisas. Quando duas pessoas conversam, não se sabe de antemão o rumo que a conversa tomará. Quando se inicia um trabalho de pesquisa, as descobertas que surgem impactam os passos subsequentes. Quando se planeja uma viagem, mesmo o mais detalhado dos roteiros consegue apenas dar um breve contorno daquilo que virá de fato a acontecer, com todos os seus pormenores. São apenas alguns exemplos para ilustrar.

Ao oferecer uma proposta para a questão do dualismo mente-corpo, os enativistas trazem o corpo para o centro para que a partir dele se possa compreender a mente. Não à toa, a origem do enativismo remonta ao já citado *The Embodied Mind* (*A Mente Corporificada*, sem versão para o português) – de Varela, Thompson e Rosch, publicado em 1991 com segunda edição de 2016. Stewart, Gapenne e Di Paolo (2010) destacam a explosão recente de pesquisas decorrentes do emergente paradigma enativista. Eles explicam que, nessa perspectiva, a cognição está “baseada na dinâmica sensório-motora das interações entre um organismo vivo e o seu ambiente” e indicam que a percepção adquire um novo posicionamento, o qual caracterizam como *radicalmente diferente*. Isso porque, no esquema clássico das abordagens cognitivas, a percepção havia sido relegada a um módulo preliminar baseado unicamente em *inputs* sensoriais a serem seguidos por uma sequência linear de “planejamento” cognitivo, a formação de representações de ações como metas/objetivos e, como ponto culminante, a decisão para agir. No esquema tradicional, a cognição estaria “sanduíchada” entre duas camadas, a do *input* sensorial e a do *output* motor, nenhum dos quais considerado propriamente cognitivo.

A metáfora do “sanduíche” para caracterizar esse ciclo supostamente linear de percepção-e-ação, com a cognição no meio como recheio, foi proposta por Hurley (1998). Essa era a ideia predominante nas abordagens cognitivistas e representacionalistas para a mente humana. Cada vez mais, no entanto, esses dois “pães do sanduíche” têm sido identificados como partes ligadas, codependentes e codeterminantes (Di Paolo, Buhrmann & Barandiaran, 2017), o que significa uma superação da ideia de que a percepção seria o início de toda atividade cognitiva, sendo a ação e a percepção duas ações sobrepostas e que acontecem

simultaneamente. *Enação* vem de *colocar em ação*, bem como de *agir enquanto percebe e de perceber enquanto age*; essas funções cognitivas, assim, se tornam concomitantes e inseparáveis. Tal ideia está presente de modo geral nas abordagens enativistas, que apresentam variações. A seguir, esclareço brevemente qual a perspectiva do enativismo que adoto nesta tese, uma vez que há algumas vertentes que vêm se desenvolvendo em direções sensivelmente diferentes dentro desse quadro.

4.2

Sobre a vertente do enativismo aqui adotada

A literatura em ciência cognitiva e na filosofia da mente e da cognição indica a existência não de um enativismo, ou de uma abordagem enativista para a cognição, mas de enativismos, por assim dizer. É comum, por exemplo, que obras ou artigos de autores enativistas comecem com uma espécie de preâmbulo procurando definir o que está sendo compreendido como esse paradigma. Esse é o caso de *The Feeling Body* (Colombetti, 2014), em que a autora faz uma leitura da afetividade na cognição a partir do enativismo. Na introdução – mencionando Thompson, Stewart, Gapenne e Di Paolo como continuadores das ideias de Varela, Thompson e Rosch – ela levanta a questão:

O que, no entanto, é o enativismo? Há várias vertentes que chamam a si mesmas de “enativismo”. A versão que utilizo neste livro é aquela que foi originalmente articulada por Varela, Thompson e Rosch (1991) (...) Por causa de sua complexidade, não é fácil caracterizar a abordagem enativa de maneira sucinta, e sempre se corre o risco de dar pouco valor a alguns aspectos importantes. Com o objetivo de fazer uma introdução, eu destaco somente alguns dos principais temas e ideias enativistas (...) Eu não endereço todas as facetas da abordagem enativa neste livro. (Colombetti, 2014, p. xiv)⁹⁰

Em *Sensorimotor Life – An Enactive Proposal*, obra na qual os autores desenvolvem aquela que chamam de “enactive theory of sensorimotor agency” (Di Paolo, Buhrmann & Barandiaran, 2017, p. 7), eles classificam a teoria sobre experiência perceptual desenvolvida por Alva Noë e Kevin O’Regan como

⁹⁰ No original: *What, however, is enactivism? There are now various accounts that call themselves “enactive”. The version of enactivism on which I draw in this book is the one originally articulated by Varela, Thompson, and Rosch (1991) (...) Because of its complexity, it is not easy to characterize the enactive approach briefly, and one always runs the risk of overlooking some important features. For the purposes of this introduction, I outline only some of enactivism’s major ideas and themes (...) I do not address all facets of enactivism in this book. (Colombetti, 2014, p. xiv)*

amplamente conhecida e elaborada, além de muito debatida, e explicitam que é “altamente compatível com muito daquilo que se quer dizer sobre agência sensório-motora”. Ressaltam, porém, que pretendem avançar para desenvolver uma interpretação não representacional de tal teoria. A questão das representações, como apontam os autores, é controversa e ambígua entre os proponentes da teoria sensório-motora para a percepção⁹¹. De todo modo, Di Paolo e parceiros concordam com parte da teoria de Noë e colegas, reconhecendo aspectos os quais trazem para as suas próprias elaborações.

Também Gallagher e Bower (2014) apontam para diferenças entre vertentes do enativismo em um trabalho cujo foco é a ênfase na afetividade e na interação social⁹². Definindo o enativismo como “uma versão das abordagens corporificadas à cognição que vêm sendo recentemente desenvolvidas” (Gallagher & Bower, 2014, p. 232), indicam que se trata de uma perspectiva mais informada pela fenomenologia e pelo pragmatismo do que outras vertentes da cognição corporificada, e que tem entrado em debates com essas outras abordagens. Essas argumentações, eles enfatizam, passam pelas questões envolvendo o funcionalismo, o papel do corpo, a maneira como nos acoplamos ao ambiente, entre outras. Ainda assim, não se pode identificar o enativismo de uma maneira generalizada, sob um mesmo rótulo, já que

[h]á várias diferenças que podem ser encontradas dentro do próprio campo do enativismo, de modo que o enativismo “inicial” de Varela, Thompson e Rosch (1991), continuado por De Jaegher e Di Paolo (2007) de maneira bastante próxima, não é exatamente o mesmo do enativismo “do meio” de O’Regan e Noë (2011; NOË, 2004) ou do enativismo “último” de Hutto e Myin (2013). Há algumas diferenças óbvias entre esses autores em termos de *pedigree*: Varela et al. são fortemente influenciados pela fenomenologia (assim como o budismo), O’Regan e Noë pela ciência cognitiva, e Hutto e Myin pela filosofia da mente analítica. (Gallagher & Bower, 2014, p. 232-233)⁹³.

⁹¹ Representações são um terreno vasto e pantanoso de estudo dentro do campo das ciências cognitivas e da filosofia da mente de modo geral. Por questão de recorte, o tema não é aprofundado neste trabalho, havendo apenas algumas referências quando elas se fazem necessárias para a compreensão de determinados pontos.

⁹² Gallagher, S. A. and Bower, M. (2014). Making enactivism even more embodied. *Avant*, 5 (2), 232-247.

⁹³ No original: *At the same time, however, there are a number of differences to be found within the enactive camp itself, so that the “early” enactivism of Varela, Thompson and Rosch (1991), closely continued by De Jaegher and Di Paolo (2007), is not exactly the same as the “middle” enactivism of O’Regan and Noë (2001; Noë, 2004), or the “latest enactivism” of Hutto and Myin (2013). There are some obvious differences among these authors simply in terms of pedigree: Varela et al. are strongly influenced by phenomenology (as well as Buddhism), O’Regan and Noë by cognitive*

Outra contribuição importante de Gallagher e Bower (2014, p. 233) é a constatação de que, apesar do foco da abordagem na percepção – que se volta para as contingências sensório-motoras – o enativismo vai além: propondo uma abordagem mais geral da cognição, inclui julgamentos perceptivos e processos de cognição superior como deliberação, tomada de decisão, memória e outros.

Não tenho o intuito de explorar profundamente as diferenças entre enativismos aqui, ainda que por vezes se faça necessário que me refira a elas com o objetivo de complementar uma determinada elaboração ou informação. Considerarei importante pontuar, de forma ligeira, esses diferentes caminhos do enativismo apenas para deixar claro que os conceitos apresentados neste capítulo, e com os quais trabalho na tese e em especial no capítulo seguinte a este, foram elaborados por uma dessas vertentes, não podendo ser tomados como todo e *qualquer* enativismo. Devo então deixar claro que, em minha análise, utilizo os conceitos tais como propostos por Di Paolo, De Jaegher e outros, especialmente em Di Paolo, Buhrmann e Barandiaran (2017); Di Paolo, Cuffari e De Jaegher (2018); Di Paolo, Rohde e De Jaegher (2010). As elaborações de Colombetti (2014) especificamente para a afetividade sob uma perspectiva enativista são igualmente consideradas. Apesar de a autora deixar claro que assume o enativismo tal como elaborado originalmente por Varela, Thompson e Rosch, suas contribuições são incluídas por conta de seu trabalho ser especificamente voltado para a afetividade. Os também citados Thompson (2010) e Maturana e Varela (2019) têm diversas ideias alinhadas com os supramencionados enativistas, sendo também incluídos pontualmente.

Nas seções subsequentes, procuro apresentar a lógica do pensamento enativista a partir de pontos que considero centrais. Começo pela questão da delimitação da unidade do ser cognoscente, um aspecto particularmente pertinente ao enativismo que não apenas é um dos que distinguem essa abordagem de outras como permite que sejam compreendidos os conceitos de *autonomia*, *adaptatividade* e *sense-making*.

science, and Hutto and Myin by analytic philosophy of mind. (Gallagher & Bower, 2014, p. 232-233)

4.3

A identidade do ser cognoscente: onde começa a cognição?

Como já mencionado, para além do cognitivismo, têm despontado na ciência cognitiva abordagens que procuram conectar diversos elementos em busca de compreender a mente para além do cérebro. Refiro-me ao esquema teórico dos *4Es*, conjunto do qual fazem parte as abordagens corporificadas a que Gallagher e Bower se referem e que citei na seção acima. A abordagem situada (*embedded*) tem como foco o enraizamento do sujeito, considerando a importância do ambiente cultural, histórico, material e social para que ele se desenvolva cognitivamente. Trata-se de uma vertente de grande importância por libertar o sujeito de uma perspectiva fechada em si mesmo, que considera seu aparato cognitivo interno suficiente para a sua operação mental. A abordagem corporificada (*embodied*) propõe que a mente inclui o corpo de maneira constitutiva, pelo menos nas suas teses (ou versões) mais fortes. A mente, assim, não apenas deve ser *explicada* por meio das conexões entre o cérebro e todas as partes do corpo, como é *formada* por um amálgama entre essas partes. Já a abordagem estendida (*extended*) inclui na atividade cognitiva aspectos externos ao organismo, considerando como partes constitutivas da cognição elementos do mundo – físico, histórico, social, cultural e tecnológico.

Apesar de considerarem essencial trazer elementos para além do cérebro para compreender a mente, penso que parte dessas abordagens mantêm em comum com o cognitivismo o seguinte elemento fundante: adotando uma perspectiva em terceira pessoa, isto é, de observação externa do sujeito, elas têm como ponto de partida uma busca por aquilo de que o cognoscente, essencialmente centrado no seu cérebro, *precisa* para agir cognitivamente. Isto é, apesar de se encarregarem de discutir o que possivelmente constitui a mente e a cognição humana, não põem em pauta o que ou quem poderia ser esse sujeito conhecedor, ou o que ou quem poderia exercer essa atividade de conhecer o mundo. O que procuro expressar aqui é que essas abordagens não fazem um esforço no sentido de delimitar o ser cognoscente; escusam-se de abordar a constituição da *identidade* desse ser e como essa identidade pode ser mantida. Em linha com tal lógica, essas perspectivas, então, tomam para si a concepção, *a priori*, de que existe o ser que conhece. Consequentemente, existe algo interior e exterior a ele, de modo que no exterior precisa haver algo que atenda a aquilo que se encontra dentro do ser. Compreendo

que, assim, corre-se o risco de meramente situar, corporificar e estender o próprio sujeito que *pensa, logo existe* proposto por Descartes (2012).

Sem dúvida, porém, é preciso admitir que, ao menos nas versões mais fortes das teses dos 4Es, procura-se superar o cartesianismo justamente a partir da premissa de que o corpo faz parte da mente, bem como o ambiente. Na cognição estendida tal como proposta por Clark (2008, 2014; ver também Clark & Chalmers, 1998), os constantes *loopings* ou trocas com o ambiente, as interações com seus elementos e as tecnologias compõem as mais diversas camadas da cognição de um ser. Descobrimos o mundo como um desenhista ou artista plástico “descobre” a sua obra por meio dos inúmeros esboços que faz até que fique satisfeito; trata-se não de um mundo que está pronto à espera de ser desnudado, mas de algo que ganhacorpo, isto é, que surge justamente a partir dessa contínua atividade sobreposta de tentativa e erro. Os elementos do mundo com os quais esboçamos e geramos esse mundo são parte constitutiva da nossa mente. Somos também as tecnologias, por isso somos *ciborgues natos* (Clark, 2003; Clark, 2014). Se, na perspectiva da cognição estendida, o ambiente é considerado constitutivo da mente, na cognição situada, de modo geral, ainda se concebe o ambiente como um apoio (Rowlands, 2014). No caso da cognição corporificada, é tênue a linha entre conceber o corpo como constitutivo da cognição ou como uma parte importante dela, mas ainda coadjuvante do cérebro.

Foge ao escopo desta tese discutir as variadas vertentes dos 4Es; no entanto, considere fundamental pontuar esses aspectos para reforçar a minha argumentação e a minha escolha teórica pelo enativismo dos supracitados autores. Como mencionei no começo desta seção, percebo que o enativismo⁹⁴ adota um ponto de partida diferente dos demais Es quando busca a compreensão da cognição humana, deslocando o ponto de partida um pouco mais para trás. Isto é, em vez de simplesmente admitir a existência de um sujeito – que conhece, que tem um cérebro, um corpo e vive no mundo – e procurar compreender um conjunto de necessidades (internas) que ele tem e que poderiam levar às suas atividades cognitivas, com o objetivo básico de resolver essas necessidades a partir da sua atividade no mundo, a perspectiva procura, antes de tudo, uma compreensão daquilo que o sujeito cognoscente é e de como ele se individua em relação ao

⁹⁴ Considerando aqui a abordagem da maneira como defini anteriormente (que a partir de agora chamo simplesmente de enativismo).

restante do mundo. E se esforça para fazer isso a partir do ponto de vista da primeira pessoa, isto é, do próprio cognoscente. Vale ressaltar que a proposta do conceito de *sense-making* amplia a possibilidade de existência de cognição (em algum nível) mesmo que não haja um sistema nervoso atuante. O enativismo procura, ainda, investigar de que maneira essa relação com o mundo pode ser constitutiva do próprio ser. Abordarei mais profundamente alguns desses aspectos posteriormente.

Sendo assim, para essa abordagem, o sujeito que conhece não está “dado”, mas configura, em si, uma importante pergunta de pesquisa. Consequentemente, no lugar de partir do princípio de que há esse sujeito e de que ele é delimitado por parâmetros pré-fixados, indo assim direto para a compreensão de como ele “consome” o mundo a sua volta, ou o utiliza para resolver seus problemas, o enativismo procura investigar que *tipo de ser* poderia ser considerado um cognoscente e como ele se relaciona com o mundo em que se insere, para além de utilitarismos e de uma relação apenas de absorção desse ambiente (privilegia-se uma relação de trocas, de mutualidade). Fundamentada na biologia, essa análise leva a um empenho por delimitar o que seria o corpo desse ser cognoscente para então investigar como se dão essas trocas com o mundo – universo o qual, como veremos, ele habita e constrói ao mesmo tempo. É por conta disso, aliás, que no vocabulário enativista há uma revisão daquilo que significa ser um “agente”.

“O foco de toda abordagem científica para agência é oferecer um enquadramento naturalizado daquilo que significa, para um sistema, agir por si próprio. Esta frase precisa ser desencaixotada” (Di Paolo, Buhrmann & Barandiaran, 2017, p. 111)⁹⁵. Os autores, com isso, querem dizer que é necessário que sejamos mais específicos ao definir o que seria agência para as ciências da mente. Um dos motivos pelos quais essa necessidade se apresenta é porque é necessário ter mais precisão quanto aos critérios que utilizamos para distinguir um sistema das redes de processos em que ele está emaranhado, ou para explicar o que pode significar exatamente um sistema agir em seu próprio interesse, ou até mesmo para justificar quais os tipos de processos que constituem um agente ou o ambiente. Procurando, então, ir além da ideia geral, Di Paolo, Buhrmann e Barandiaran (2017, p. 111) assinalam três requisitos mínimos para que um sistema seja considerado um

⁹⁵ No original: *The aim of any scientific approach to agency is to provide a naturalized account of what it means for a system to act on its own behalf. This phrase needs to be unpacked.* (Di Paolo, Buhrmann & Barandiaran, 2017, p. 111).

agente, dentro do contexto das ciências da mente. São os seguintes: 1) *autoindivuação* (*self-individuation*), 2) *assimetria interacional* (*interactional assymetry*) e 3) *normatividade* (*normativity*). A célula seria um exemplo de sistema mínimo a contar com os três itens. É também a unidade mínima da vida: a mente, para o enativismo, começa onde a vida começa.

Em relação ao primeiro requisito, segundo a abordagem enativa, agentes são sistemas que, de maneira ativa, são capazes de demarcar a sua individuação. “Somente sistemas capazes de se autossustentar e de se distinguir de seu entorno, e que enquanto o fazem definem um ambiente em que sua atividade se desdobra, são candidatos a agentes nesta abordagem” (Di Paolo, Buhrmann & Barandiaran, 2017, p. 112). Esta definição será melhor clarificada adiante, quando for abordado o conceito de *clausura operacional*.

O segundo requisito, *assimetria interacional*, refere-se às regulações que o agente é capaz de fazer quando se relaciona com o ambiente. As trocas com esse ambiente precisam ser equilibradas, e então o acoplamento do sistema com o ambiente é, nesse sentido, simétrico – mesmo que haja diferenças entre os fluxos de energia e matéria do agente para o ambiente e do ambiente para o agente, há uma relação simétrica de codependência entre esses sistemas. “No caso particular de sistemas que se autoindivuem, como sistemas vivos ou autopoieticos, a noção de *acoplamento estrutural* é também simétrica nesse sentido”⁹⁶ (Maturana & Varela, 1980, 1987, apud Di Paolo, Buhrmann & Barandiaran, 2017, p. 116). A assimetria, por outro lado, refere-se às maneiras como o agente pode modular esses fluxos, influenciando o ambiente. Por vezes, o agente pode alterar os parâmetros ou as condições para o seu acoplamento com o ambiente.

O terceiro requisito é o da *normatividade*. Trata-se de um conceito justificado pela natureza do próprio agente. Com o conceito de autoindivuação em mente, certas modulações assimétricas desempenhadas pelo sistema dão suporte aos processos que o distinguem de seu meio; enquanto isso, outras modulações podem interferir com esses processos, ameaçando quebrar o sistema. O que define quais movimentos caem em qual categoria?

⁹⁶ No original: *In the particular case of self-individuating systems, such as living or autopoietic systems, the notion of structural coupling is also symmetrical in this sense (Maturana & Varela, 1980, 1987).*

Pelo menos uma fonte de normas intrínseca (...) se origina da própria organização do sistema que o mantém como uma entidade que se auto distingue e se produz a si mesma. Nesse sentido, as ações do agente e os eventos do ambiente podem ser bons ou maus para a sua continuada existência (Di Paolo, Buhrmann & Barandiaran, 2017, p. 122).

As regulações que o sistema é capaz de empreender associam-se ao conceito de *adaptatividade*, que significa a capacidade do agente de ativamente monitorar e regular seus processos a partir de normas intrínsecas. Adaptar-se é ser capaz de melhorar as condições de vida e evitar ou enfrentar ameaças à sua viabilidade, lidando com uma situação a partir das normas inerentes à autoindivuação do agente. Nessas situações, o agente deve agir dentro de gradações e de uma maneira direcionada. A adaptatividade é uma pré-condição para o *sense-making* (Di Paolo, Buhrmann & Barandiaran, 2017, p. 122-123). Nem sempre o organismo consegue agir da melhor forma para adaptar-se, mas ele é orientado a isso.

É interessante notar que mesmo nestes casos o comportamento produzido ainda segue as normas no sentido de que o resultado de uma ação que falha só pode ser discernido de um movimento neutro no contexto da norma implícita nos processos adaptativos envolvidos. (Di Paolo, Buhrmann & Barandiaran, 2017, p. 123)⁹⁷

Essas ideias voltarão a aparecer no decorrer do trabalho. Na seção a seguir, introduzo alguns outros conceitos e ideias que são fundamentais para a compreensão da proposta do enativismo: a *autonomia*, a noção de *sense-making*, a maneira como os enativistas propõem a corporificação, os conceitos enativistas de emergência e de experiência e as ideias de individuação e *becoming*. Dentre esses conceitos e ideias, o de **autonomia** tem destaque na argumentação que elaboro no capítulo a seguir. Entender como os enativistas propõem a autonomia é essencial para a compreensão do conceito de *sense-making* e de *participatory sense-making* que, junto à ideia enativista de *experiência*, explicitam elementos cruciais acerca da cognição e, conseqüentemente, da aprendizagem humana. Esses elementos constituem parâmetros para a caracterização de seres humanos e de sistemas artificiais. Por mais que essa diferenciação não seja o objetivo central da minha argumentação, considerá-la faz parte do trabalho de pesquisa aqui envolvido, já que

⁹⁷ No original: *It is interesting to note that even in such cases the behavior produced still follows norms in the sense that the outcome of an action that fails can only be distinguished from a neutral movement in the context of the norm implied by the adaptive processes involved. (Di Paolo, Buhrmann & Barandiaran, 2017, p. 123).*

procurar entender essas diferenças pode nos ajudar a compreender a aprendizagem humana e também como os sistemas artificiais podem fazer parte do nosso cotidiano de aprendizagem contínua, corporificada e intersubjetiva.

4.4 Autonomia

Tal como proposto pelos autores em cujo trabalho fundamento esta tese, o enativismo surge a partir de uma associação extensiva e mútua entre a vida e a mente. Isto significa que, com uma forte raiz na biologia, procura conceber a mente pelas mesmas vias pelas quais se pode conceber a vida. Busca especificar como se poderia definir um ser vivo e como ele se delimita e se individua. Na articulação de conceitos enativistas fundamentais, fica claro que, enquanto a associação com o computador parecia “cair como uma luva” para uma perspectiva da mente humana que concentrava no cérebro todo o processamento cognitivo – a mente como um software rodando em um hardware cerebral – uma outra imagem emerge a partir da perspectiva enativista: a da mente como uma célula viva, autônoma. Sua membrana garante a organização, a individuação e a identidade do meio interno em relação ao externo, simultaneamente garantindo as trocas essenciais entre os dois meios, numa relação de codependência.

A abordagem enativa oferece uma tentativa de fundamentar a cognição usando categorias que descrevem a organização dos sistemas vivos. Em vez de entender a cognição como um processo semelhante ao de um computador, o enativismo começa por considerá-la um processo semelhante à vida ancorado no corpo vivo. Em vez de ser apenas um substituto da mente, um recipiente, veículo ou instrumento controlado pelo cérebro, mesmo que capaz de implementar funções computacionais inteligentes, o corpo – vivo e vivido, o material e o corpo experiencial – é concebido como a fonte de tudo o que a mente é e pode ser. (Di Paolo, Buhrmann & Barandiaran, 2017, p. 20)⁹⁸

O paradigma da enação resolve este problema, tomando a cognição como uma característica essencial dos organismos vivos. Para Descartes, não houve problema em considerar que todos os animais são meras máquinas, então a divisão ontológica viria entre animais e humanos. Para Maturana e Varela (1980) e Jonas (1963), pelo

⁹⁸ No original: *[T]he enactive approach provides an attempt to ground cognition using categories that describe the organization of living systems. Instead of understanding cognition as a computer-like process, enactivism starts by considering it as a lifelike process anchored in the the living body. Instead of being merely a surrogate of the mind, a container, vehicle, or instrument controlled by the brain, even one that could implement clever computational functions, the body – the living and the lived body, the material and the the experiencing body – is conceived as the source of all that the mind is and can be. (Di Paolo, Buhrmann & Barandiaran, 2017, p. 20)*

contrário, a grande divisão vem entre a matéria e os organismos vivos. (Stewart, 2010, p. 1)⁹⁹

Para além da metáfora, quando Maturana e Varela – precursores do enativismo que vem sendo desenvolvido por Di Paolo, De Jaegher e colegas – apresentaram a célula como unidade básica da vida, eles desejavam responder à questão: “como saber quando um ser é vivo?” (Maturana & Varela, 1984: 2019, p. 48). A proposta desses dois biólogos foi definir os seres vivos como aqueles que apresentam uma organização autopoietica, conceituada como a capacidade de produzirem a si próprios de modo contínuo. Segundo Maturana e Varela (1984: 2019, p. 52-54), esse tipo de organização demanda que os elementos componentes do ser estejam relacionados numa rede contínua de interações delimitada por uma membrana. Mas, ao mesmo tempo em que delimita a célula, circunscrevendo a sua identidade, a membrana é parte constitutiva dessa rede de relações e transformações entre seus componentes. Isto é,

[p]or um lado, é possível perceber uma rede de transformações dinâmicas, que produz seus próprios componentes e é a condição de possibilidade de uma fronteira; de outra parte vemos uma fronteira, que a é a condição de possibilidade para a operação da rede de transformações que a produziu como uma unidade (Maturana & Varela, 1984: 2019, p. 54).

É preciso que uma série de elementos se mantenham em equilíbrio para que a célula se mantenha viva e, assim, siga individuada. Ela depende tanto das trocas com o meio como de sua atividade individual, de sua operação e de sua organização. Uma célula cuja membrana se desfaz deixa de ser uma célula, desaguando no ambiente, assim perdendo-se, desfazendo-se; deixando de *ser*.

O conceito de autopoiese descreve um aspecto peculiar da organização dos organismos vivos, a saber, que seus processos contínuos de trocas materiais e energéticas com o mundo, e de transformação interna e metabolização, se relacionam de tal forma que a mesma organização é constantemente regenerada pelas atividades dos próprios processos, apesar das variações que ocorrem de caso para caso. (Di Paolo & Thompson, 2014, p. 69)¹⁰⁰

⁹⁹ No original: *The paradigm of enaction solves this problem by grounding all cognition as an essential feature of living organisms. For Descartes, there was no problem in considering that all animals are mere machines, so the ontological split came between animals and humans. For Maturana and Varela (1980) and Jonas (1963), by contrast, the great divide comes between matter and living organisms.* (Stewart, 2010, p. 1)

¹⁰⁰ No original: *The concept of autopoiesis describes a peculiar aspect of the organization of living organisms, namely that their ongoing processes of material and energetic exchanges with the world, and of internal transformation and metabolizing, relate to each other in such a way that the same*

Sendo a individuação um dos aspectos essenciais endereçados pelo enativismo, a autonomia surge como um conceito que dá suporte a essa concepção. Afinal, para que um organismo seja autônomo, ele deve ter uma identidade bem demarcada que, ao menos em algum nível, seja autogerada – o que o caracteriza como uma entidade distinta. Como diz a citação acima de Di Paolo e Thompson, esse sistema deve ser capaz, simultaneamente, de se regenerar e se adaptar. “Organismos vivos são autônomos – eles seguem normas estabelecidas por sua própria atividade” (Di Paolo, Rohde & De Jaegher, 2010, p. 37; Di Paolo, 2005)¹⁰¹. Isto significa que o cognoscente não pode estar sujeito estritamente a leis externas a ele, pois, se assim for, não é capaz de se orientar em prol de si mesmo, de traçar o curso de sua vida e de sua história.

Um sistema cuja identidade é totalmente especificada por um designer e não pode, por meio de suas próprias ações, regenerar sua própria constituição, só pode seguir as leis contidas em seu projeto, não importa o quão plástico, adaptável ou vivo seja o seu desempenho. Para que um sistema possa gerar suas próprias leis, ele deve ser capaz de construir a si mesmo em algum nível de identidade. Se um sistema “não tem voz” na definição de sua própria organização, então está condenado a seguir um desenho que lhe foi atribuído externamente, como uma via férrea (Di Paolo, Rohde & De Jaegher, 2010, p. 35)¹⁰².

Como indicam Di Paolo e Thompson (2014, p. 69), o conceito de autonomia responde por aquele que mais fortemente demarca a diferenciação entre o enativismo e outras abordagens corporificadas à cognição. Ele pode ser tomado como uma ideia que está enraizada na teoria da autopoiese tal como proposta por Maturana e Varela – tanto é que o enativismo que Di Paolo e colegas desenvolvem, fundamentado sobre a noção de autonomia, tem sido identificado por pesquisadores do campo como “enativismo autopoietico” (Hutto & Myin, 2013). Porém, há variações importantes na forma como os enativistas se apropriam da teoria da autopoiese que Maturana e Varela desenvolveram – o que, segundo Di Paolo e

organization is constantly regenerated by the activities of the processes themselves, despite whatever variations occur from case to case. (Di Paolo & Thompson, 2014, p. 69)

¹⁰¹ No original: *Living organisms are autonomous — they follow laws set up by their own activity. (Di Paolo, Rohde & De Jaegher, 2010, p. 37; Di Paolo, 2005)*

¹⁰² No original: *A system whose identity is fully specified by a designer and cannot, by means of its own actions, regenerate its own constitution, can only follow the laws contained in its design, no matter how plastic, adaptive, or lifelike its performance. In order for a system to generate its own laws, it must be able to build itself at some level of identity. If a system “has no say” in defining its own organization, then it is condemned to follow an externally given design like a railroad track (Di Paolo, Rohde & De Jaegher, 2010, p. 35).*

Thompson (2014, p. 69), implica em que esse rótulo oferecido à abordagem enativista corra o risco de não ser tão preciso.

A autopoiese seria o mecanismo por meio do qual os seres vivos se caracterizam como autônomos. Thompson (2007, p. 37) – que destaca ideias relacionadas à teoria dos sistemas dinâmicos que, segundo ele, são basilares para a abordagem enativista – indica que um sistema autônomo é um sistema que se autodetermina. Pode ser uma célula, um animal multicelular ou uma colônia, como mencionado; ou um ser humano; todos se comportam de uma maneira coerente, como uma unidade capaz de determinar a si mesma em suas interações com o ambiente. De acordo com Varela (1979), a autonomia foi concebida, inicialmente, como uma generalização do conceito de autopoiese ou autoprodução. Mais tarde foi que Varela estendeu o conceito a outros domínios, depois de identificar uma lógica de organização similar em sistemas nervosos animais e redes imunitárias. O conceito veio a ser aplicado, ainda, aos campos da comunicação (redes comunicacionais) e da conversação (Di Paolo & Thompson, 2014).

A ideia de autoconstituição, no enativismo, conecta-se com a ideia de *clausura operacional* (*operational closure*) e de *precariedade* (o sistema autônomo é um sistema operacionalmente fechado – *operationally closed* – e precário – *a precarious system*). Como explicitado em Di Paolo (2009, p. 16; 2021) e Colombetti (2014, p. 15-16)¹⁰³, trata-se de um tipo de organização na qual os resultados das operações desempenhadas pelos elementos constitutivos do sistema se mantêm dentro do próprio sistema. Porém, a realização física da autonomia, como eles destacam, depende de que essa rede – cuja operação é fechada – seja termodinamicamente aberta, isto é, capaz de manter trocas de matéria e energia com o meio para regular a sua atividade. Por isso, a ideia de “operacionalmente fechado” não implica em um isolamento energético ou material, mas em um modo específico de relacionamento funcional entre os componentes do sistema.

O crucial, portanto, é o seguinte: ainda que o agente autônomo demarque a sua individualidade, constituindo-se a si mesmo de uma maneira contínua, sua atividade cognitiva está baseada nas trocas constantes entre o cognoscente e o ambiente em que ele se encontra imerso. Isto é, autonomia não implica em independência do meio; pelo contrário. A vida do agente tem com o ambiente uma

¹⁰³ Ver também Di Paolo, Rohde e De Jaegher (2010, p. 38).

relação de *needful freedom*, como Hans Jonas (1966) colocaria; trata-se de uma relação de codependência, orientada ao ambiente. Na medida em que age no ambiente e se movimenta nele, o agente é afetado por esse meio não de maneira passiva, meramente respondendo a “perturbações externas”, mas, “de fato, ativamente e de forma assimétrica, *regula* as condições de suas trocas com o ambiente, e ao fazer isso *põe em ação* (*enact*) um mundo ou domínio cognitivo” (Di Paolo, Rohde & De Jaegher, 2010, p. 38)¹⁰⁴. Todo sistema autônomo tem um propósito (Colombetti, 2014, p.16). Pode-se dizer que o ser sobrevive *apesar* das instabilidades do meio, ou justamente *por causa* delas: afinal, é graças às instabilidades ou perturbações (chamadas pelos enativistas de *condições precárias*) a que se encontra constantemente submetido que o cognoscente aprimora a sua capacidade de automanutenção, garante a sua autonomia e mantém a sua organização.

Um ponto fundamental é que, para além de ser autopoietico e autônomo, para ser considerado um ser cognoscente no sentido enativista é preciso ter capacidade *adaptativa*, um aspecto já citado como essencial para caracterizar um agente. A adaptatividade é uma das responsáveis por demarcar as diferenças sensíveis entre o conceito original de autopoiese e o de autonomia enativista. “Di Paolo argumenta que *sense-making* demanda adaptatividade, mas a autopoiese mínima não é suficiente para a adaptatividade, e então é insuficiente para o *sense-making*”¹⁰⁵ (Thompson, 2009, p. 148). O conceito de *sense-making* é trabalhado no tópico que se segue a este, em que a questão da adaptatividade também será retomada.

Outro ponto essencial da autonomia enativista é que o equilíbrio entre o agente cognitivo e o meio em que ele se insere depende de trocas equilibradas com esse ambiente. Existem as normas que regem esse intercâmbio e mantêm o seu equilíbrio, e é na atividade de conhecer o mundo que o ser se apropria dessas normas, ao mesmo tempo em que contribui para concebê-las (isso tem relação direta com *sense-making*). Por conta disso, e de alguns outros fatores, Di Paolo e Thompson ressaltam que, para eles, diferente do que Varela postulou, autonomia e clausura operacional não têm o mesmo significado.

¹⁰⁴ No original: [*In fact actively and asymmetrically regulate the conditions of their exchange with the environment, and in doing so, enact a world or cognitive domain. (Di Paolo, Rohde & De Jaegher, 2010, p. 38).*]

¹⁰⁵ No original: *Di Paolo argues that sense-making requires adaptivity, but minimal autopoiesis is not sufficient for adaptivity and hence is insufficient for sense-making. (Thompson, 2009, p. 148).*

Autonomia é o mesmo que fechamento operacional? Embora pareça ter sido o mesmo para Varela (1979) – pelo menos em termos de sua definição formal de um sistema autônomo – argumentamos em outros escritos que o fechamento operacional não é suficiente para capturar certas implicações importantes do sentido mais amplo do termo "autonomia" - implicações expressas por noções como "espontaneidade", "intencionalidade imanente", "teleologia intrínseca" e a "autogeração de normas" (ver Di Paolo, 2005; Thompson, 2007). Dada a definição acima de fechamento operacional, vários exemplos triviais de tal fechamento podem existir. Por exemplo, nos autômatos celulares, a regeneração de um estado de equilíbrio em cada célula depende mutuamente do estado de equilíbrio em outras, transformando as dependências em uma rede fechada. (Di Paolo & Thompson, 2014, p. 71-72)¹⁰⁶.

O que torna o conceito de clausura operacional um dos pilares da autonomia, mas não equivalente ao conceito como um todo, é a ideia de *precariedade*. Mais uma vez, enativistas propõem para esse termo não o sentido trivial ou do senso comum, mas um significado específico. Esse significado tem a ver com as intrincadas relações que se estabelecem numa rede operacionalmente fechada.

Quando um processo é ativado pela rede operacionalmente fechada e também por processos externos, se a rede é removida e o processo permanece - nas novas circunstâncias - graças apenas ao apoio externo, então esse processo não é precário. Um processo precário é tal que, qualquer que seja a complexa configuração das condições de operação (da rede), se as dependências de outros processos na rede operacionalmente fechada forem removidas, o processo necessariamente pára. Em outras palavras, não é possível que um processo precário em uma rede operacionalmente fechada exista por si só nas circunstâncias criadas pela ausência da rede. (Di Paolo & Thompson, 2014, p. 72)¹⁰⁷

Como esclarece Stapleton (2022, no prelo), mesmo para uma bactéria o mundo é precário porque, se a bactéria se encontrar numa situação em que não haja fornecimento suficiente de nutrientes de que precisa para viver, e não fizer nada quanto a isso, sua organização irá se desintegrar. “Ela precisa de energia para gerar

¹⁰⁶ No original: *Is autonomy the same as operational closure? Although it seems to have been the same for Varela (1979)—at least in terms of his formal definition of an autonomous system—we have argued in other writings that operational closure is not sufficient to capture certain important implications of the wider sense of the term “autonomy”—implications conveyed by notions such as “spontaneity,” “immanent purposiveness,” “intrinsic teleology,” and the “self-generation of norms” (see Di Paolo, 2005; Thompson, 2007). Given the above definition of operational closure, various trivial examples of such closure may exist. For example, in cellular automata, the regeneration of an equilibrium state in each cell mutually depends on the equilibrium state in others, making the dependencies into a closed network. (Di Paolo & Thompson, 2014, p. 71-72)*

¹⁰⁷ No original: *When a process is enabled by the operationally closed network and by external processes as well, if the network is removed and the process remains—in the new circumstances—thanks only to the external support, then that process is not precarious. A precarious process is such that, whatever the complex configuration of enabling conditions, if the dependencies on other processes in the operationally closed network are removed, the process necessarily stops. In other words, it is not possible for a precarious process in an operationally closed network to exist on its own in the circumstances created by the absence of the network. (Di Paolo & Thompson, 2014, p. 72)*

os processos dinâmicos constantes que produzem as partes da célula que delimitam a sua membrana” (Stapleton, 2022, no prelo, p. 165). Sem essa demarcação, que é o que a individua, a bactéria deixa de ser um organismo.

4.5 Emergência

O conceito de *emergência* é usado para descrever a formação de uma nova propriedade ou de um novo processo a partir da interação de processos ou eventos diferentes já existentes antes dele. Como consequência do acoplamento do ser com o mundo, algo novo *emerge* que não existiria sem que esse acoplamento acontecesse.

A autonomia não é propriedade de uma coleção de elementos, mas consequência de uma nova identidade que surge de processos dinâmicos em clausura operacional precária. O significado não se encontra nos elementos pertencentes ao ambiente ou na dinâmica interna do agente, mas pertence ao domínio relacional estabelecido entre os dois (Di Paolo, Rohde & De Jaegher, 2010, p. 40)¹⁰⁸.

Pode-se distinguir um processo emergente de uma simples agregação de elementos dinâmicos a partir de dois fatores (Di Paolo, Rohde & De Jaegher, 2010, p. 40):

- 1) O processo autônomo deve ter a sua identidade, autônoma também;
- 2) A sustentação dessa identidade e a interação entre o processo emergente e seu contexto deve levar a modulações do funcionamento dos níveis subjacentes.

O enativismo (Di Paolo, Rohde & De Jaegher, 2010, p. 40-42) indica que essa interação entre níveis deve ser endereçada caso a caso, por vezes recorrendo-se a conceitos e ferramentas complexos derivados da teoria de sistemas dinâmicos. Os autores indicam que o cognitivismo, por sua vez, teria associado níveis pessoais e subpessoais de maneira indiscriminada, explicando propriedades de níveis

¹⁰⁸ No original: *Autonomy is not a property of a collection of elements, but the consequence of a new identity that arises out of dynamical processes in precarious operational closure. Meaning is not to be found in elements belonging to the environment or in the internal dynamics of the agent, but belongs to the relational domain established between the two (Di Paolo, Rohde & De Jaegher, 2010, p. 40).*

superiores a partir de níveis inferiores, como se já estivessem “magicamente presentes” desde sempre.

O primeiro exemplo de emergência, no contexto do enativismo e correspondente ao sentido que essa abordagem atribui ao conceito, é o da vida em si. Autopoiese é o processo de autoprodução contínua, de renovação, de regeneração de uma rede física de transformações moleculares; esse processo não é encontrado em nenhum nível abaixo daquele da célula viva. O conceito de emergência proposto pelos enativistas provoca uma mudança importante na maneira de abordar a questão sobre como uma determinada função cognitiva acontece. A partir da perspectiva do enativismo, não faz sentido definir propriedades a partir de componentes isolados. Em vez de nos perguntarmos qual a localização ou a função de um componente cognitivo em específico, devemos questionar quais interações precisam ocorrer para dar origem a determinada função cognitiva.

Emergência é um dos pilares do conceito de *sense-making* – que apresento na próxima seção – e de outro que é basilar do enativismo. Trata-se do conceito de *participatory sense-making* ou *sense-making participativo*, o qual será tratado em mais detalhes no capítulo seguinte a este.

4.6 ***Sense-making***

O enativismo encapsula no conceito de *sense-making* boa parte das ideias acerca daquilo que concebe como cognição. Um organismo autônomo, um agente, é capaz de fazer sentido do/para o mundo em que vive, que habita e que co-constrói. Se isso parece já muito claro, por outro lado pode ir contra alguns pressupostos tradicionalmente aceitos. Conforme citado, cognição tomada como *sense-making* não demanda que exista um sistema nervoso “que represente fatos do mundo, raciocínio sobre eles e gere regras para a ação” (Colombetti, p. 18). Pela lógica enativista, mesmo um organismo muito simples pode produzir sentido do/para o seu mundo. Afinal, sua atividade é guiada por elementos que, estando a sua volta, representam algo para a sua sobrevivência. *Sense-making*, assim, torna-se um conceito desafiador, indo além daquilo que mais tradicionalmente se compreende como fazer sentido do mundo.

Este é um conceito que tem a ver com o fato de que seres vivos se deixam *afetar* pelo meio; isto é, o ambiente, o contexto material, histórico, social e cultural, as relações, o espaço, nunca podem lhes ser indiferentes. Variados mundos de significado e de sentido podem emergir para diferentes seres, de acordo com as suas características corporais, orgânicas, intersubjetivas, sensório-motoras, as suas percepções, a sua história e diversos outros fatores. Para o cognoscente, dentro da lógica enativista, o mundo que existe a sua volta passa a fazer sentido quando ele é sensibilizado pelos elementos do entorno de alguma maneira. Isto é, na medida em que é afetado pelo meio, o cognoscente inclui em seu repertório esses diversos aspectos, características e componentes do mundo, que assim passam a formar sua rede de relações e significados.

Sendo assim, a compreensão da construção de sentido e de significado, pelas lentes do enativismo, exige que se adote uma perspectiva bastante específica. Desafia-se a ideia de que existe informação pronta para ser consumida ou absorvida pelos organismos, como o cognitivismo implica.

Organismos não recebem informação passivamente de seu ambiente para depois transformá-las em representações. Sistemas cognitivos naturais simplesmente não se prestam a acessar o mundo para construir imagens dele. Eles participam da geração de significado/sentido por meio de seus corpos e de sua ação, frequentemente se engajando em interações transformacionais e não meramente informativas; *they enact a world*. (Di Paolo, Rohde & De Jaegher, 2010, p. 39).¹⁰⁹

Nesse contexto, o mundo não é uma ilusão, mas somente passa a constituir o espaço do cognoscente quando é reconhecido e compreendido por ele. O enativismo, ao menos a vertente que teve origem com Varela, Thompson e Rosch (1991/2016), procura um meio termo entre o realismo e o idealismo, como indicam Rolla e Figueiredo (2021). São considerados, nessa perspectiva, o realismo como a vertente filosófica que postula que nós acessamos um mundo que já existe, por meio de representações mentais; e o idealismo como a linha que defende que a cognição resulta da projeção dos nossos próprios processos cognitivos, relacionando-se exclusivamente com as representações internas do agente, que é desacoplado do ambiente.

¹⁰⁹ No original: Organisms do not passively receive information from their environments, which they then translate into internal representations. Natural cognitive systems are simply not in the business of accessing their world in order to build accurate pictures of it. They participate in the generation of meaning through their bodies and action often engaging in transformational and not merely informational interactions; *they enact a world*. (Di Paolo, Rohde & De Jaegher, 2010, p. 39).

Rolla e Figueiredo (2021) indicam que o dilema com o qual o enativismo lida não está totalmente solucionado e aparece de maneiras diferentes em vertentes variadas do enativismo. Apontam, ainda, para a ênfase garantida pelos enativistas mais ao agente do que ao ambiente, nessa que é uma perspectiva que propõe o acoplamento entre os dois. Em busca de um caminho para a compreensão de onde o enativismo, tal como desenvolvido por Di Paolo e colegas, se encontra nesse dilema realismo-idealismo, os autores argumentam que a proposta do enativismo de *enact* ou *bring forth a world* deve ser considerada de forma literal; isto é, há um mundo exterior a nós, com suas estruturas físicas externas, mas somos nós que o fazemos existir. Esse mundo é o resultado da nossa atividade cognitiva; agente e mundo determinam um ao outro – isso envolve a evolução e inclui o corpo, o cérebro e o ambiente acoplados, acontecendo no decorrer do tempo, em diferentes escalas de tempo, curtas ou extensas (Rolla & Figueiredo, 2021, p. 7). Esta é uma das formas de *enact* ou *bring forth a world* que esses dois autores propõem. Eles apontam outras duas.

A segunda forma seria a construção de nichos em prol da nossa sobrevivência. Ao empreendermos esses esforços, eles retornam a nós e modificam quem somos e o que podemos fazer. Essa perspectiva evoca uma concepção da adaptação dos organismos ao meio, não somente como resultado das pressões que recebem do exterior e que desafiam a sua permanência, mas também como esforços que empreendem no sentido de tornar o meio mais adaptado a sua sobrevivência, provocando mudanças ecológicas. As transformações no meio se tornam, assim, uma espécie de herança não genética que deixam para as gerações subsequentes, “inaugurando novos caminhos de evolução” (Rolla & Figueiredo, 2021, p. 8).

A construção de nichos sociais seria a terceira forma de *bring forth a world* apontada pelos autores e estaria em continuidade com as outras duas. Nós, humanos, não herdamos apenas genes e transformações ambientais promovidas pelos nossos antepassados, mas artefatos, informações sobre como utilizá-los e sobre como podemos evoluí-los; herdamos, ainda, normas e instituições; e é a combinação de elementos dessa natureza, que orientam e delimitam as fronteiras para o nosso comportamento, que pode ser considerada uma terceira maneira de *fazer emergir um mundo*.

Percebemos um martelo não só como algo que pode ser segurado, mas como uma ferramenta, ou uma ferramenta de alguém, talvez uma ferramenta valiosa ou barata;

percebemos os livros não apenas como algo que podemos folhear, mas como legível, agradável ou enfadonho; percebemos sinais de alerta não apenas como objetos com os quais podemos colidir, mas como proibindo-nos de seguir um caminho ou fazer certas coisas em uma área específica e assim por diante. (Rolla & Figueiredo, 2021, p. 9)¹¹⁰

Creio que uma imagem que ajuda a esclarecer essa ideia enativista é a do míope de grau muito alto que usa óculos de lentes poderosas¹¹¹; sem as lentes, pode ver apenas borrões em torno de si, um emaranhado de cores e formas desfocadas, sem linhas, sem limites, sem contorno. Quando coloca os óculos no rosto, consegue enxergar as formas bem delineadas e começa a compreender o que elas lhe “dizem”. Isto é, existe um mundo que está ali, mas ele não pode ser acessado por completo sem alguma interação cognitiva. Outro exemplo seria o seguinte: imagine um texto todo escrito em um determinado idioma que você não compreende. Em determinada parte do texto, há uma frase, essa sim, num idioma que você conhece. Essa frase acaba “saltando aos olhos”, por ser numa língua conhecida e, conseqüentemente, compreensível. O conhecimento de um idioma se dá meio de uma construção em que diversos elementos entram em jogo; esse processo demanda interação e intersubjetividade. Desse modo, ainda que os escritos estejam lá, somente são decifrados e passam efetivamente a fazer parte do mundo de um agente se ele conseguir compreendê-los, e isso se liga não apenas à história pessoal e individual dele, mas a uma sociedade, já que aquilo foi construído coletivamente. As interações cognitivas e coletivas são, enfim, os elementos que tornam o meio de fato um meio para nós, no qual vivemos; os exemplos de Rolla e Figueiredo (2021), a seguir, são, também, bastante elucidativos.

Nada nos proíbe de descrever, para diversos fins científicos, um mundo existente independentemente de qualquer atividade e experiência humana. Por exemplo, podemos descrever a Terra como um objeto astronômico sujeito apenas a forças físicas, sem nenhuma menção explícita à humanidade. E, na medida em que estas descrições sejam comprovadamente frutíferas, podemos aceitá-las como verdadeiras. A Terra concebida como um corpo astronômico é independente da mente nesse sentido, mas não é o mundo com o qual nos envolvemos diretamente, não é um mundo vivido. E, apesar de sua independência da mente, ela não é pré-determinada – afinal, ela é descrita através das leis de uma disciplina feita pelo

¹¹⁰ No original: *[w]e perceive a hammer not only as graspable but as a tool, or someone's tool, maybe a valuable or cheap one; we perceive books not Only as something we can flip through, but as readable, enjoyable or boring; we perceive warning signs not only as objects with which we can collide, but as forbidding us from following a path or doing certain things in a specific area and so on.*

¹¹¹ Exemplo autobiográfico; sou míope de 7 graus num olho e 5 no outro.

homem, que é cumulativamente trazida por vários antes de nós (Rolla & Figueiredo, 2021, p. 10)¹¹²

Considere outro exemplo astronômico: a primeira vez que a equipe da missão espacial Apollo 11 chegou à lua. Os astronautas de fato fizeram emergir um novo ambiente por meio de sua chegada à lua – o que, obviamente, não significa dizer que eles construíram um corpo celeste. Mas o fato de que a lua, desde então, tornou-se disponível para uma investigação direta, e as subseqüentes consequências sociais e políticas da corrida ao espaço mostram que a lua se tornou parte do nosso mundo vivido de uma nova forma. (Rolla & Figueiredo, 2021, p. 10)¹¹³

À semelhança desse processo, pode-se dizer que *sense-making* remete a uma rede de sentidos e significados que vai sendo construída na medida em que o cognoscente habita, constrói, entranha-se no mundo a que pertence, tornando-o de fato o seu espaço, o seu ambiente, que é também o espaço de outros que vieram antes dele e que virão depois. Não é possível fazer surgir o mundo de outra forma que não seja essa, que envolve tateá-lo, testá-lo, experimentá-lo e, sucessivas vezes, estabelecer com ele uma relação de estranhamento, dúvida, questionamento; é instigando, vivendo e investigando que o cognoscente vai revelando a si mesmo esse universo a que pertence e que lhe é indissociável. Até o mesmo o próprio corpo pode ser visto como algo que somente passa a ser acessível a nós por completo na medida em que nos familiarizamos com ele; afinal, precisamos conhecer nossos limites e possibilidades (orgânicas, sensoriais e motoras e intersubjetivas, dimensões que serão apresentadas mais adiante). Enativistas referem-se ao corpo que viveu experiências como *living body* (Di Paolo & Thompson, 2014).

Trazendo novamente a concepção básica enativista de que percepção e ação ocorrem juntas, pode-se constatar que, ao mesmo tempo em que o cognoscente percebe o mundo que habita e que ajuda a construir, ele atribui sentido, significado aos seus elementos na medida em que os “toca”, os tangibiliza, e assim os inclui em sua rede extensiva de componentes que ajudam a garantir sua organização e

¹¹² No original: *Nothing forbids us from describing, for various scientific purposes, a world existing independently of any human activity and experience. For instance, we can describe the Earth as an astronomical object subject only to physical forces, with no explicit mention of humanity at all. And, insofar as these descriptions are proven fruitful, we can accept them as true. Earth conceived as an astronomical body is mind-independent in that sense, but it is not the world we directly engage with, it is not a lived world. And despite its mind-independence, it is not pre-given—after all, it is described through the laws of a human made discipline, which is cumulatively brought forth by many before us.*

¹¹³ No original: *Consider another astronomical example: the first Moon landing by the crew of the Apollo 11 mission. The astronauts did bring forth a new environment by landing on the Moon—which obviously is not to say that they have built a celestial body. But the fact that the Moon is since then available to direct, hands-on investigation and the socio-political consequences of the space race show that the Moon became part of our lived world in a new way.*

autonomia. É por isso que os enativistas destacam que a produção de sentido está intimamente relacionada à noção de *Unwelt*, cunhada por Uexküll ([1934] (2010)). "Para um sistema vivo, ser um sistema capaz de *sense-making* é viver em um mundo que é sempre um *Unwelt*, ou seja, um ambiente que tem um significado ou valor específico para ele" (Colombetti, 2014, p. 17)¹¹⁴. Para que esse mundo nítido, esse mundo que tem um sentido, emergja para o cognoscente – como aos olhos do espectador da arte surge uma certa imagem diante de uma pintura impressionista – é indispensável o acoplamento entre esse ser e o mundo¹¹⁵. Isto é, a ação do cognoscente é fundamental, mas ela precisa se dar nesse mesmo mundo que está ali e que se transforma no mundo dela a partir da sua ação. “Não é que o significado/sentido seja encontrado em elementos que pertencem ao ambiente ou às dinâmicas internas do agente; ele pertence ao domínio relacional estabelecido entre os dois” (Di Paolo, Rosch & De Jaegher, 2010, p. 40).

4.6.1

Uma inversão relevante na compreensão tradicional da cognição

Não somente a ideia de *sense-making* é essencial como conceituação da cognição dentro da proposta enativista, mas traz na sua essência uma diretriz fundamental para entendermos aquela que pode ser considerada uma inversão importante promovida pela abordagem. Comumente, se concebe a *cognição superior* como raciocínio, memória, tomada de decisão e tudo aquilo que parece mais complexo e complicado; conseqüentemente, associamos essas capacidades cognitivas à atribuição de significado, à capacidade de dar sentido ao mundo. Segundo essa lógica, os seres cognoscentes capazes de tal feito seriam aqueles dotados de habilidades cognitivas mais elevadas, habitando assim o “topo” dessa “cadeia”. Para entender as capacidades humanas, procura-se compreender como elas funcionam em termos de mecanismos mais simples que podem ou não estar presentes em animais não humanos; por isso diz-se que o cognitivismo é uma abordagem “*top down*” (Stapleton, 2022, no prelo).

No entanto, se considerarmos *sense-making* como uma capacidade que está não no topo, mas na base da cognição, podendo ser estendida a diversos seres –

¹¹⁴ Essa noção se aproxima, ainda, do conceito de *affordances*, proposto por James Gibson (1986).

¹¹⁵ Há, na arte, uma concepção que contribui para ilustrar esta comparação que faço aqui; trata-se da ideia de *beholder's share* (ver Seth, 2017).

mesmo a aqueles cujos organismos são aparentemente mais simples – passamos a vê-la como uma capacidade primordial: a cognição começa pelo *sense-making*. Concebe-se então que, a partir dessa capacidade intrínseca a organismos vivos acoplados ao meio, outras capacidades sejam desenvolvidas. Pode-se pensar numa espiral que, na medida em que ganha corpo e se complexifica, adquire novos contornos; nem todos os seres têm as mesmas capacidades – e o enativismo não indica isso, pelo contrário, respeita o que irá emergir a partir da ação do ser em seu mundo em variados níveis.

Abordagens estendidas e enativas podem então ser pensadas como “bottom up”, já que o desafio para os seus pesquisadores é explorar como essas habilidades ‘escalam’ até o nível da cognição humana: o quanto da cognição humana consegue ser explicado sem se ter que recorrer a computações internas e representações? (Stapleton, 2022, no prelo).¹¹⁶

Parece que essa inversão – isto é, sentido e significado surgindo na base em vez do topo da pirâmide – não apenas distancia o enativismo do cognitivismo (e de outros aspectos de outros dos 3Es) como pode ajudar a explicar aquilo que menciono no começo deste capítulo. Isto é, as tarefas que supostamente seriam mais fáceis para robôs fazerem, mas que continuam sendo as mais difíceis e intrigantes para esses sistemas artificiais, podem ser tão desafiadoras justamente porque talvez não sejam de fato as mais simples. Possivelmente são das mais sofisticadas, por serem intrinsecamente associadas justamente à condição de ser um organismo vivo, autônomo, um corpo que habita um meio que de certa maneira é, em si, parte constitutiva desse ser. E isso indica, ainda, que olhar para as habilidades que estão na base da cognição pode explicar as capacidades ditas “superiores”. Para ilustrar essa ideia, abaixo apresento esta figura que é uma adaptação da original, presente em Di Paolo, Rohde e De Jaegher (figura 2.1, cuja legenda é “*Life-mind continuity and the scale of increasing autonomy and mediacy*”; consta em Di Paolo, Rohde e De Jaegher, 2010, p. 50). Na imagem original, as setas são descendentes e o conceito de autopoiese e teleologia, bem como os demais, sucessivamente,

¹¹⁶ No original: *Extended and enactive approaches might thus be thought of as “bottom-up” approaches as the challenge for these researchers is to then explore how well these kind of perceptual and cognitive abilities “scale up” to human-level cognition: How much of human cognition can be explained without having to appeal to internal computations over representations?* (Stapleton, 2022, no prelo)

apareciam no topo da imagem. Considerei oportuno “virar” a figura, para que fique clara a interpretação que aqui defendo.

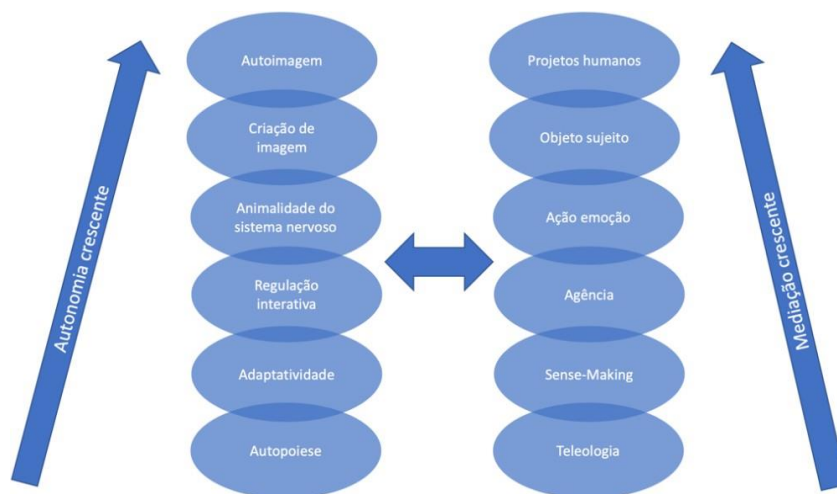


Figura 2 – Continuidade vida-mente (adaptada de Di Paolo, Rohde & De Jaegher, 2010, p. 50)

No primeiro nível da pirâmide estão os organismos que somente são capazes de reagir a perturbações do entorno, sem monitorar de maneira lógica os seus estados e maximizar as condições para a sua autopoiese continuada. É justamente o monitoramento que permite isso. Em contraste, nota-se o conceito de *sense-making* logo acima da base da pirâmide. Ao lado do conceito de *sense-making* aparece a *adaptatividade*, que, como citei, é um elemento essencial a diferenciar autonomia de mera autopoiese (é necessário ter a capacidade de se adaptar para ser considerado autônomo no sentido enativista). Para retomar essas ideias, cito Rolla (2020), que explica que a adaptatividade é a capacidade que possibilita a movimentação de um sistema vivo em direção a aquilo que é favorável a ele e para longe daquilo que lhe pode ser prejudicial:

A autopoiese sozinha não é suficiente para explicar o comportamento antecipativo e flexível de sistemas vivos (Di Paolo, 2005). Isso ocorre porque a autopoiese explica como sistemas vivos aproximam-se de certos estados ambientais e afastam-se de outros tendo em vista exclusivamente a conservação da sua organização, isto é, mantendo-se dentro dos seus limites de viabilidade. Contudo, existe uma gradação ou um contínuo entre estados *mais ou menos favoráveis* e estados *mais ou menos prejudiciais* para um sistema em um ambiente suficientemente complexo (como frequentemente é o caso). Além disso, um sistema vivo deve ser capaz de reconhecer estados ambientais *possivelmente* favoráveis ou *possivelmente* prejudiciais para que efetive ajustes estruturais anteriormente ao encontro com esses estados – seja para garantir a aproximação a estados favoráveis ou para garantir o distanciamento de estados prejudiciais. (Rolla, 2020, p. 52)

Quanto maior a capacidade de mediação, maior a autonomia do cognoscente. Com o aumento da mediação, a tendência é que o organismo seja capaz de atribuir significado de uma maneira que não é apenas guiada pelas suas necessidades metabólicas. Assim, a pirâmide tem vários níveis porque organismos de vários tipos exibem diferentes “graus de autonomia e mediação” (Di Paolo, Rohde & De Jaegher, 2010, p. 51).

Animais, através de sua atitude, demonstram capacidade de agir e perceber, bem como medo ou desejo por algo distante. E os seres humanos, através de capacidades socialmente mediadas como a de fazer imagens e, finalmente, de construir uma autoimagem, ganham a capacidade de considerar as situações objetivamente e de se definir como sujeitos, de distinguir a verdade da falsidade e de experimentar a felicidade e a frustração (Jonas, 1966; Di Paolo, 2005, apud Di Paolo, Rohde & De Jaegher, 2010)¹¹⁷.

Em resumo, pode-se dizer que a proposta enativista é a de que fazer sentido é a fonte da cognição, não uma consequência da atividade cognitiva; é nessa perspectiva que tal atividade se encontra na base, na raiz – isto é, não se trata apenas de um recurso para compreendermos a cognição, mas de fato essa é parte da ontologia proposta por essa vertente de pesquisa para a cognição. A cognição começa com esse fazer sentido do mundo. Na medida em que o cognoscente sofisticada a sua capacidade de agência, ele adquire outras capacidades cognitivas. Não à toa, a regulação interativa aparece no terceiro patamar da pirâmide, já que a intersubjetividade é um dos pontos cruciais do enativismo.

4.7

Linguistic bodies: a corporificação e a linguagem no enativismo

O enativismo e a chamada cognição corporificada compartilham algumas ideias concernentes ao corpo na cognição. No entanto, em determinadas vertentes da abordagem corporificada, o corpo segue como um apoio, não necessariamente como a fonte de onde a cognição emerge. No enativismo, como se pode notar, a cognição deve ser concebida antes de tudo *a partir do corpo* porque é dele que ela

¹¹⁷ No original: *Animals, through their motility, exhibit the capacity to act and perceive as well as fear or desire for something distal. And humans, through socially-mediated capacities such as image-making and ultimately for constructing a self-image, gain the ability to regard situations objectively and define themselves as subjects, to distinguish truth from falsehood, and to experience happiness and frustration (Jonas, 1966; Di Paolo, 2005, apud Di Paolo, Rohde & De Jaegher, 2010).*

se origina. Devido a essa centralidade, os enativistas preocupam-se em descrever que corpo é esse e como ele é composto. Em vez de um artefato com mecanismos controlados pelo cérebro, o corpo é, em si, um sistema animado em diversas camadas.

O corpo não é uma marionete controlada pelo cérebro, mas todo um sistema animado com muitas camadas autônomas de autoconstituição, autocoordenação, auto-organização e vários graus de abertura ao mundo que geram a sua atividade de fazer sentido. (Di Paolo, Rohde & De Jaegher, 2010, p. 42)¹¹⁸

O corpo proposto pelo enativismo conta com três dimensões: orgânica, sensório-motora e intersubjetiva. Segundo os enativistas, há evidência de que habilidades cognitivas de nível superior, como raciocínio e resolução de problemas, manipulação mental de imagens e uso de linguagem dependem de estruturas corporais, de maneira crucial. O ambiente é indissociável do corpo em suas três dimensões, as quais se sobrepõem e se complementam mutuamente, pois é a partir das trocas constantes com o meio externo que o ser conhece e ajuda a moldar o mundo que habita. É também justamente porque está inserido no mundo e nessa rede de trocas incessantes que o ser se individua, sendo esse acoplamento indispensável para a demarcação da sua individualidade e identidade.

No entanto, para além de serem constituídos individualmente por um conjunto de tronco, cabeça e membros, corpos são peças de uma engrenagem maior; nesse encaixe, cada sujeito tem não um corpo, mas corpos, e esses corpos se vinculam a outros, de outros indivíduos. Nossos corpos trazem marcas não apenas da nossa história pessoal, mas daquela que envolve os nossos antepassados, os lugares pelos quais passaram, os períodos em que viveram, as pessoas com quem se envolveram. Segundo os enativistas, não é possível compreender a cognição humana sem entender que ela se constitui por meio dessa interconexão entre mentes de seres vivos em constante transformação. Somos *linguistic bodies*, ou corpos linguísticos. E vivemos um processo contínuo de *becoming* – tornar-se (Di Paolo, 2020). Estamos, sempre, nos transformando, mesmo quando parecemos exatamente os mesmos.

¹¹⁸ No original: *The body is not a puppet controlled by the brain but a whole animate system with many autonomous layers of self-constitution, self-coordination, and self-organization and varying degrees of openness to the world that create its sense-making activity* (Di Paolo, Rohde & De Jaegher, 2010, p. 42)

Não é difícil perceber o quanto modulamos, regulamos ou ajustamos as nossas ações, pensamentos e até os nossos sentimentos a partir dos movimentos dos nossos corpos. Por vezes, o corpo parece tomar uma decisão antes de pensarmos sobre ela. Isso acontece quando nos vemos, por exemplo, acuados ou pressionados em determinada situação e então agimos “por impulso” – assim definimos frequentemente as ações que parecem ter sido espontâneas, “do nada”. Mas, segundo a linha de raciocínio enativista, nenhuma ação é “do nada”. É como se aquela decisão fosse gestada no corpo do agente, até emergir. Também não podemos separar decisões aparentemente tomadas pelos nossos corpos de decisões tomadas pelas nossas cabeças. Eles são inseparáveis. Uma decisão “protagonizada” pelo corpo não tem menos valor do que uma decisão planejada; e, mesmo uma decisão planejada, no sentido de ser pensada previamente, envolve todas as dimensões da nossa corporificação. Além disso, se tomamos decisões e voltamos atrás, quantas vezes não é justamente porque *sentimos* algo que nos faz voltar – algo que, por vezes, a princípio é inexplicável?

Essa não linearidade, essa convivência mútua de avanços e retrocessos, por assim dizer, é uma característica da mente humana. Também notamos esse vaivém quando analisamos a sociedade, a humanidade: diversos movimentos sociais, culturais e políticos convivem, bem como crenças, religiões, rituais, legislações; por exemplo, a própria questão sobre direitos humanos serem ou não universais ainda gera dificuldades em grupos que acentuam diferenças que, segundo alegam, justificam o estabelecimento de direitos e deveres diferentes. Observamos, por exemplo, o avanço e a negação da ciência em simultâneo, e a inteligência artificial explodindo em suas diversas facetas, seja contribuindo para prever impactos climáticos ou para agravá-los, para ajudar as pessoas a encontrarem empregos ou para dificultar essa busca. Não à toa, enativistas propõem que temos três dimensões ou camadas corporais que convivem, já mencionadas: a orgânica, a sensório-motora e a intersubjetiva. Cada uma dessas camadas tem em seu seio uma forma de autonomia, já que cada uma garante um dos ciclos de funcionamento do corpo. Os ciclos de regulação do organismo referentes ao corpo todo e as sensações de dor, fadiga, sede, fome etc. pertencem à dimensão orgânica; essa é a camada da relação do corpo com o mundo em termos de necessidades e cuidados (Di Paolo, Buhrmann & Barandiaran, 2017, p. 5). Nessa dimensão localiza-se a autonomia metabólica, fisiológica.

Os ciclos de acoplamento sensório e motor entre o organismo e o ambiente pertencem à dimensão sensório-motora. Ainda segundo Di Paolo, Buhrmann e Barandiaran (2010, p. 5-6), trata-se daquela camada em que os elementos orgânicos se somam a essas outras articulações; na medida em que o dia se desenrola, sentimos o peso do cansaço, ou a satisfação de cumprir determinadas tarefas, por exemplo. Somos permeados por esses sentimentos, humores, estados de espírito. Nesta dimensão, a autonomia refere-se à clausura sensorial e motora, aos *loops* (ciclos) de ação-percepção e aos esquemas sensório-motores que se apoiam mutuamente.

A última camada da corporificação, a da interação intersubjetiva, refere-se aos ciclos de interações sociais, “sinergias transindividuais” e relações sociais (Di Paolo, Rohde & De Jaegher, 2018, p. 87). Nenhum corpo pode ser concebido sozinho, mas sim por

[u]m processo e uma perspectiva relacional que nos leva a conceber os seus envolvimento com outros corpos e com o mundo não como algo que vem depois de serem constituídos, mas como parte da sua constituição. (Di Paolo, Cuffari & De Jaegher, 2018, p. 99) ¹¹⁹.

Apesar de a dimensão intersubjetiva reforçar que, para o enativismo, a interconexão entre seres vivos é essencial para que eles existam, não algo opcional ou derivado, os autores (2018, p. 87) admitem a existência de uma certa progressão nas dimensões, nesta ordem – orgânica, sensório-motora e intersubjetiva. Isto é, a primeira precisa existir para que a segunda exista, e as duas anteriores precisam existir para a terceira existir. No entanto, reforçando a não linearidade, os autores buscam evidências empíricas que mostram como essas dimensões se afetam mutuamente e se interpenetram; um exemplo é quando alguém se esquece de comer porque passou por um episódio de raiva, envolveu-se em uma situação altamente emocional com outra pessoa; ou quando fuma ou come doces em excesso mesmo tendo a consciência de que esses hábitos podem não ser bons para a manutenção do corpo e, conseqüentemente, a autonomia no nível orgânico.

Os enativistas ressaltam a existência dessas três camadas corporais porque se preocupam em definir de que corpo(s) falam; estão trabalhando para desenvolver

¹¹⁹ No original: *[F]rom a process and relational perspective that steers us into conceiving their engagements with other bodies and with the world not as something that comes after they are constituted but as part of their constitution.*

uma teoria sobre corpos linguísticos. “Corpos são inacabados, sempre por se tornar. Em particular, corpos linguísticos” (Di Paolo, Cuffari & De Jaegher, 2018, p. 7). Em consonância com esses corpos linguísticos, faz sentido uma abordagem constitutiva para a linguagem:

Vamos perseguir uma abordagem constitutiva para a linguagem, concebendo-a não simplesmente como um conjunto de competências sofisticadas para a comunicação, expressão etc., mas como a definição de humanismo. Paralelamente a esta visão constitutiva, adotamos a perspectiva de que a linguagem não é uma entidade estática, manifestada como um conjunto de regras, vocabulários, nem mesmo uma série de eventos de comunicação. A linguagem é um fluxo vivo de atividade no mundo sócio material das práticas e da história. A linguagem é um campo de luta, transformação, crítica, de enação humana. (Di Paolo, Cuffari & De Jaegher, 2018, p. 7)¹²⁰

O já citado conceito de *becoming*, somado aos conceitos de *emergência* e de *experiência* no enativismo, ajudam na compreensão da proposta de colocar não somente o corpo, mas corpos no centro da atividade cognitiva. A ideia de experiência no enativismo pode ser tomada como um conceito mais amplo, capaz de englobar outros. Para experimentar, é preciso ter um corpo vivo e autônomo; fazer *sense-making*; relacionar-se com o mundo e com os seres desse mundo. “A experiência deve ser explicada para além dos dados” (Di Paolo, Rohde & De Jaegher, 2010, p. 43)¹²¹.

No trecho a seguir, usando o exemplo de um conhecedor de vinhos, Di Paolo, Rohde e De Jaegher destacam a diferença entre obter um certo tipo de conhecimento por meio do recebimento de *informações* ou por meio de uma *transformação*:

Tornar-se um conhecedor de vinhos é certamente um objetivo alcançável, mas a perícia neste campo (como em qualquer outro) não é obtida através da obtenção do tipo certo de informação, mas através do tipo certo de **transformação** – uma transformação que só pode ser conseguida por meio de uma formação apropriada e estendida no tempo (experimentar, cometer erros, etc.). A experiência é alterada de uma forma regular ao longo do processo. É ela própria um aspecto habilidoso da atividade corporificada. (Di Paolo, Rohde & De Jaegher, 2010, p. 44, grifo meu)¹²²

¹²⁰ No original: *We will pursue a constitutive approach to language, conceiving it not simply as a set of sophisticated skills for communication, expression, and so on, but as defining of humanness. Alongside this constitutive view, we adopt the perspective that language is not a static entity, manifested as a set of rules, vocabularies, nor even a series of communication events. Language is a living stream of activity in the social material world of practices and history. Language is a field of struggle, transformation, criticism, of human enaction.* (Di Paolo, Cuffari & De Jaegher, 2018, p. 7)

¹²¹ No original: *Experience goes beyond data to be explained*” (Di Paolo, Rohde & De Jaegher, 2010, p. 43).

¹²² No original: *Becoming a wine connoisseur is certainly an achievable goal but expertise in this field (as in any other) is not obtained through gaining the right kind of information but through the right kind of transformation — one that can only be brought about by appropriate time-extended*

Resgatando a fenomenologia, *experiência* no enativismo pode se traduzir como tudo aquilo que emerge da atividade corporal de dar sentido ao mundo. É algo que se aproxima do conceito de *skillful coping* proposto por Dreyfus (2014).

Muitos relatos modernos de atividade cognitiva já levam a experiência a sério. Por exemplo, a defesa de Dreyfus da aquisição de habilidades não representacionais é baseada na atenção cuidadosa à experiência de passar por um processo de aperfeiçoamento no desempenho de tarefas. À medida que fazemos a transição de iniciantes a especialistas na prática, não apenas o desempenho habilidoso (*skillful*) é melhorado, mas a experiência também é transformada. Isto é de se esperar se a corporificação, no sentido enativo, for levada a sério. Se a experiência e a interação corporal se relacionassem como dois sistemas mutuamente externos, esperaríamos ou uma relação imutável ou uma relação fugaz entre nossos corpos e nossa experiência. Em vez disso, encontramos uma relação legítima de transformação corporal e de experiência (Di Paolo, Rohde & De Jaegher, 2010, p. 44)¹²³.

Essas transformações corporais não podem existir isoladas dos corpos de outros indivíduos. *Becoming*, a ideia enativista de *tornar-se* ou *vir-a-ser*, remete a impressões que vão ficando em nós enquanto vivemos e experimentamos as mais diversas situações; impressões não no sentido de percepções do senso comum, mas de cicatrizes orgânicas, físico-químicas, sensoriais, motoras e também provocadas pelas interconexões com os outros. Essas marcas não apenas nos afetam, elas são aquilo que somos e que nos tornamos na medida em que nos desenvolvemos, tomamos decisões, criamos e seguimos caminhos, estruturamos modos de ver e pensar o mundo e daí por diante. Di Paolo (2021b) questiona:

Devemos seguir o esquema tradicional e pensar nos corpos principalmente a partir de uma perspectiva orgânica, anatômica, fisiológica (a forma como eles se apresentam quando são colocados sobre uma mesa de operação)? Deveríamos vê-los como constituídos por suas práticas, sua história, seu modo de vida, suas escolhas, suas relações socioambientais (e eu realmente quero dizer constituídos como em fabricados, organizados, compostos, por esses processos)? (Di Paolo, 2021b, p. 784)¹²⁴.

training (experimenting, making mistakes, and so on). Experience is altered in a lawful manner through the process. It is itself a skillful aspect of embodied activity. (Di Paolo, Rohde & De Jaegher, 2010, p. 44)

¹²³ No original: *Many modern accounts of cognitive activity already take experience seriously. For instance, Dreyfus's defense of nonrepresentational skill acquisition is based on paying careful attention to the experience of undergoing a process of task improvement. As we make the journey from beginners to experts to practice, not only is skillful performance improved, but experience is also transformed. This is to be expected if embodiment in the enactive sense is taken seriously. If experience and the body-in-interaction were to relate to each other as two mutually external systems, we would expect either an unchangeable or a fleeting relation between our bodies and our experience. Instead we find a lawful relation of bodily and experience transformations (Di Paolo, Rohde & De Jaegher, 2010, p. 44).*

¹²⁴ No original: *Should we go along with the traditional picture and think of bodies mainly from an organic, anatomical, physiological perspective (the way they look laid down on the operating*

Depois de explicitar que é na segunda opção que ele e outros enativistas investem, o filósofo complementa:

A ideia é olhar além da constituição orgânica dos corpos biológicos (autopoiese), mas mantendo-se dentro da lógica da individuação, ou seja, processos autônomos precários de constituição material e relacional ocorrendo não apenas no domínio das reações bioquímicas, mas também na organização de padrões sensoriomotores (corpos sensoriomotores) e agências sociolinguísticas (corpos linguísticos). A dimensão orgânica dos corpos não desaparece ou se desvanece quando se considera outras formas de autoconstituição corporal. Pelo contrário, de acordo com a visão enativa, as diferentes dimensões formam um fluxo enredado de processos em múltiplas escalas, um *becoming* contínuo cuja estrutura não se esgota pelos ciclos do metabolismo ou pela natureza do desenvolvimento dependente do caminho. (Di Paolo, 2021b, p. 784)¹²⁵.

Como corpos linguísticos, a linguagem é, para nós, não apenas descritiva, mas constitutiva, como coloca Bannell (2022, no prelo), citando Taylor, cujas ideias se alinham com a proposta enativista para a linguagem:

Para entender essa diferença, pense no caso do cientista que está tentando encontrar uma maneira de descrever um fenômeno melhor que outros cientistas, ou dar um nome para uma nova espécie de animal. Ele tenta achar a melhor maneira de descrever o fenômeno ou nomear a espécie e, quando consegue, ele sente que a descrição é a correta. Mas a melhor descrição ou nome não muda o objeto. Agora, pense em outro caso onde a dificuldade não é em nomear ou descrever algo que já existe independente da descrição ou nome. Pelo contrário, é a tentativa de articular algo que sentimos vagamente ou que nos ameaça ou impede nossos propósitos etc. Esses significados são sentidos. Aqui entra a dimensão afetiva. O sentido ou a emoção é, muitas vezes, uma maneira de ser consciente do objeto intencional. Esse sentido é o que Taylor chama de “intuição sentida” (*felt intuition*). (Bannell, 2022, no prelo)

table)? Should we see them as constituted by their practices, their history, their way of life, their choices, their socio-environmental relations (and I really mean constituted as in made, assembled, put together, by these processes)? (Di Paolo, 2021b, p. 784).

¹²⁵ No original: *The idea is to look beyond the organic constitution of biological bodies (autopoiesis) but keeping within the logic of individuation, i.e., precarious autonomous processes of material and relational constitution taking place not only in the domain of biochemical reactions but also in the organization of sensorimotor patterns (sensorimotor bodies) and socio-linguistic agencies (linguistic bodies). The organic dimension of bodies does not disappear or fade into a background when considering other forms of bodily self-constitution. On the contrary, according to the enactive view, the different dimensions form an entangled stream of processes at multiple scales, an ongoing becoming whose structure is not exhausted by the cycles of metabolism or the path-dependent nature of development. (Di Paolo, 2021b, p. 784).*

4.8

Conclusão do capítulo

Neste capítulo, procurei introduzir as ideias básicas oferecidas pelos enativistas e sua linha de pensamento, apresentando os fundamentos dessa perspectiva que aponta no horizonte como um possível novo paradigma para pensarmos a mente e a atividade cognitiva humana. Acentua-se o corpo como uma chave primordial a partir da qual a mente surge. Busca-se uma alternativa ao dualismo mente-corpo; não há mente separada do corpo, nem corpo separado da mente, e nem mesmo se deveria trabalhar a questão da cognição nesses termos. Quando se trata de investigar a maneira como o ser humano funciona, objetivando compreender sua relação com as tecnologias de inteligência artificial, os fatores biológicos propostos pelo enativismo se revelam um ponto crucial. Afinal, se a vida se mantém a partir do próprio vaivém entre o ser e o mundo, e também a mente depende desses movimentos: procura-se explicar como a cognição acontece a partir desses termos. Toma-se, assim, uma direção que distancia substancialmente humanos de máquinas – ao menos tais como estão concebidas até o momento.

No começo do capítulo, mencionei a questão de que sistemas de machine learning não têm experiência, não experimentam nada de fato porque não têm o arsenal biológico necessário para isso. Em termos enativistas, se poderia dizer que um sistema artificial não é capaz de reconhecer as valências primitivas na sua relação com o mundo, distinguindo ou identificando aquelas que seriam favoráveis à sua automanutenção e aquelas que, num outro extremo, se mostrariam prejudiciais a esse mesmo fim. E isso acontece justamente porque sistemas artificiais, pelo menos como aqueles que existem hoje, não são autônomos como seres vivos – este tema será tratado com mais profundidade no capítulo a seguir. Ou seja, é porque a cognição emerge a partir das interações biológicas, e com a finalidade de ampliar e facilitar essas relações, que um sistema artificial, por não ter interações biológicas fundamentais, se mostra capaz apenas uma simulação de processos propriamente cognitivos¹²⁶.

Mencionei ainda neste capítulo uma tensão primordial, assim identificada porque o ser que se individua é, ao mesmo tempo, um ser que necessita da sua

¹²⁶ Agradeço imensamente ao professor Giovanni Rolla não somente pelas críticas gerais durante o desenvolvimento desta pesquisa, mas pelas indicações específicas que levaram a esta importante correlação aqui apresentada.

relação com o meio para se individuar. O ser constitui-se pela sua clausura operacional, mas tem também uma abertura ao meio externo como condição para se manter vivo. Da mesma maneira que há essa tensão primordial quando falamos em clausura operacional e, ao mesmo tempo, em uma abertura ao ambiente, nós podemos dizer que há uma tensão entre os aspectos da individualidade e da intersubjetividade no enativismo. Para que se compreenda a proposta enativista em sua totalidade, deve-se entender como convivem esses dois elementos. O conceito de *participatory sense-making* é capaz de condensar essa indispensável atenção à atividade corporal intersubjetiva.

A partir da introdução do conceito de *participatory sense-making*, no próximo capítulo irei apresentar uma análise que investiga possíveis pontos de tensão que emergem quando o ambiente com o qual interagimos é aquele permeado por algoritmos. Em específico, argumento acerca daquilo que pode ocorrer com a autonomia, no sentido enativista do conceito, quando o sujeito cognoscente se encontra num ambiente predominantemente algorítmico. Tendo como base a noção de que a percepção que temos do mundo é diretamente impactada pelo ambiente e pela qualidade das trocas que mantemos com ele, o que se estende às trocas com outros seres vivos, vale avaliar o que acontece com a percepção e mesmo com a própria capacidade de manter trocas intersubjetivas quando estamos lidando com sistemas algorítmicos.

The desire to leave no potential data point uncollected with the objective of mapping out our future is incompatible with treating individuals as masters of their own lives. (Véliz, 2021)

Conforme procurei demonstrar no capítulo inicial desta tese, preconizava-se à época das máquinas analógicas de Skinner, Pressey e outros que esses sistemas seriam capazes de individualizar o ensino, respeitando o ritmo de cada aluno e ainda liberando o professor de tarefas repetitivas. Potencializadas por algoritmos, máquinas de certo modo semelhantes a essas antigas – já que suas premissas se assemelham às desses aparatos analógicos – agora amplificadas pela sua gigantesca capacidade computacional de lidar com grandes quantidades de dados (big data) prometem proporcionar ainda mais oportunidades nesse sentido. Elas se inspiram no cérebro e focam no seu desenvolvimento. Enquanto isso, cada vez mais e para além dos limites formais da educação, nós, como sociedade, construímos e habitamos ambientes permeados por algoritmos. O que isso pode significar para a nossa atividade cognitiva? Quais as possíveis implicações que podemos observar nesse processo, quando se trata de analisar os impactos dos sistemas algorítmicos de aprendizagem de máquina sobre a cognição humana?

Inseridos nesta espécie de *algoritmosfera*, somos ao mesmo tempo os fornecedores do material para a operação desses sistemas e os consumidores dessa operação, já que eles essencialmente se utilizam dos nossos dados para funcionar. Fazemos isso o tempo todo, ligados a diversos sistemas, muitos deles interligados entre si. Pode-se dizer que as máquinas de aprender, quiçá de ensinar, agora estão por toda a parte. E se, por um lado, a lógica das *máquinas que aprendem* passou a nos acompanhar na forma como estudamos, nos relacionamos nas redes sociais, pesquisamos na internet, assistimos a filmes e séries, selecionamos destinos de férias, nos informamos, enfim, desempenhamos as mais diversas tarefas, não é nos limites dos sistemas pré-definidos como plataformas de aprendizagem (adaptativa ou não) que se encontram as ferramentas tecnológicas baseadas em machine learning que possivelmente dão suporte à aprendizagem. Isso é importante para a discussão sobre tecnologias digitais na educação até mesmo porque, se inteligência artificial ainda soa como algo distante do cotidiano de tantas escolas, por outro lado

crece o acesso à internet pelos celulares ou smartphones. Em 2019, esses dispositivos representavam mais de 98% dos tipos de equipamentos usados para ter acesso à rede entre brasileiros com mais de dez anos de idade¹²⁷. Em Portugal, o acesso à internet via smartphone cresceu de 9% de penetração em 2010 para 72% em 2018, mantendo-se assim até 2020¹²⁸. Basta ter acesso à internet para que haja compartilhamento de dados, esteja o usuário ciente disso ou não. Portanto, não podemos mais falar em educação, sob a perspectiva ampla da formação humana, sem buscar compreender o que significa estarmos imersos em uma cultura baseada em dados, comumente referida como economia de dados.

Num cenário em que circula muita informação sobre inteligência artificial, big data e algoritmos, a compreensão do sentido de “learning” embutido em machine learning tende a se difundir. No entanto, mesmo que se popularize a noção de que a aprendizagem de máquina significa basicamente uma tentativa de “prever o futuro” com base em estatística, assim se relativizando o sentido de aprendizagem nesse contexto, ainda há muito espaço para dúvidas a respeito do efetivo poder desses sistemas. Por exemplo, questiona-se até onde vai sua capacidade real, ainda que saibamos que é delimitada matematicamente. Sistemas artificiais podem chegar a agir como agimos? Vão nos superar em termos de capacidades cognitivas? O que esses sistemas de fato são capazes de fazer? Em meio a tais inquietações, há um tópico subjacente: a possibilidade de autonomia das máquinas. Trata-se de um tema que se liga a questões controversas relacionadas à ética dos sistemas, à responsabilidade que pode ou deve ser atribuída a eles e aos seus processos de decisão (se assim puderem ser considerados) e daí por diante. Uma outra dimensão da questão – mais pertinente a este trabalho – se liga às dúvidas sobre as reais capacidades cognitivas, se assim puderem ser chamadas, dos sistemas artificiais assentados na aprendizagem de máquina.

A partir de tantos questionamentos, neste capítulo gostaria, então, de propor um enfoque particular para o tema da autonomia, e o faço porque considero essa perspectiva pertinente a uma discussão sobre aprendizagem sob as lentes enativistas. No lugar de focar na autonomia das máquinas, gostaria de me

¹²⁷ Fonte: <https://educa.ibge.gov.br/jovens/materias-especiais/20787-uso-de-internet-televisao-e-celular-no-brasil.html>

¹²⁸ Segundo estudo da Bareme Internet/Marktest; fonte: <https://www.marktest.com/wap/a/n/id~269e.aspx>

concentrar naquilo que possivelmente acontece com a *autonomia dos seres humanos*, particularmente enquanto agentes cognoscentes, quando lidamos com sistemas de machine learning. Parto da hipótese de que a *autonomia* do cognoscente – adotando-se aqui a perspectiva enativista – seria afetada de algum modo quando utilizamos tais sistemas. Eles constituem, afinal, ao mesmo tempo um meio composto por nós e que nos afeta e nos molda. Sendo assim, investiguei possíveis implicações para isso, as quais procuro apresentar.

Sem dúvida, admito que debater a autonomia humana na *algoritmosfera* exige olhar ao menos para alguns dos aspectos mais relevantes das possibilidades de autonomia, independência ou agência atribuída a sistemas artificiais – esses domínios, afinal, se entrecruzam. No entanto, ressalto que a minha decisão por focar na autonomia dos agentes, humanos, em vez de me aprofundar na possível autonomia dos sistemas artificiais deve-se ao fato de que esta é uma pesquisa que busca compreender a aprendizagem de máquina a partir da cognição humana, procurando entender como a segunda pode ser impactada pela primeira. Autonomia, por sua vez, é um dos conceitos centrais do enativismo, como foi apresentado no capítulo anterior. É também um conceito em constante debate na educação.

No enquadramento proposto agora, os sistemas artificiais algorítmicos de machine learning são tomados como o meio que nós, agentes cognoscentes, construímos coletivamente e coabitamos. Isto é, o ambiente digital permeado por machine learning é, simultaneamente, criado e usufruído por nós. Como já mencionado, produzimos uma massa de conteúdo, composta pelos chamados dados, e esses dados que nós produzimos são aplicados pelas redes neurais para fazer previsões¹²⁹. Sendo assim, poderíamos perguntar, tomando por base o enativismo: que tipo de mundo de significado e de sentido podemos constituir em torno de nós por meio desse processo de gerar dados e, ao mesmo tempo, de ser impactados pelos próprios dados que geramos (e seus usos)? Conseguimos manter a nossa autonomia, no sentido enativista, quando envolvidos por esses nichos

¹²⁹ *Machine learning algorithms analyze the data to identify patterns and to build a model which is then used to predict future values (for example [...] by identifying patterns in photographs of named people, it predicts who is shown in other photographs; and by identifying patterns in medical symptoms, it predicts a specific diagnosis). In other words, machine learning may be considered a three-step process (analyze data, build a model, undertake an action) that is continuously iterated (the outcomes of the action generate new data, which in turn amends the model, which in turn causes a new action). It is in this sense that the machine is learning (Holmes, Bialik & Fadel, 2019, p. 89).*

algorítmicos? Como eles nos afetam e afetam a nossa experiência como seres cognoscentes, uma vez que estamos imersos em um mundo que, para além de uma fonte de *inputs*, é uma arena para agirmos e fazermos aparecer os resultados das nossas ações e da nossa percepção? Preocupo-me, sobretudo, com a aquisição de conhecimento. Lidar com esses sistemas nos ajuda a expandir esse processo de conhecer ou acaba por reduzi-lo?

Argumento que a nossa autonomia como *sense-makers* sofre perturbações e tende a adquirir novos limites quando lidamos com sistemas de machine learning. Isso acontece porque esses sistemas não podem ser considerados *sense-makers* ou agentes autônomos e, como consequência e de maneira complementar a isso, não há realmente interação quando lidamos com eles, ou pelo menos não há interação social, uma vez que a interação (mais uma vez, no sentido enativista) é algo que acontece entre dois ou mais agentes¹³⁰ (Di Paolo, Rohde & De Jaegher, 2010). Por trás dessa argumentação encontra-se um aspecto muito pertinente ao enativismo, que é o da intersubjetividade, somado ao elemento da afetividade.

No desenvolvimento de meu argumento, além de articular os conceitos básicos do enativismo, fundamento-me no conceito de *sense-making* participativo, ou *participatory sense-making*¹³¹. O objetivo de trazer esse conceito, que constitui uma extensão do conceito de *sense-making*, é articulá-lo para buscar compreender a nossa relação com os sistemas algorítmicos. Ainda que esteja consciente quanto à existência de inúmeras questões éticas, políticas e culturais que têm emergido na sociedade algorítmica em que vivemos, como em toda esta tese a perspectiva de análise adotada é a da cognição, o foco continua sendo nos impactos sobre a aprendizagem humana, aqui compreendida desde a percepção e de modo amplo e contínuo. A dimensão da aprendizagem está conectada com as dimensões éticas, estruturais, sociais, culturais, estéticas, antropológicas; porém, a minha preocupação é em procurar contribuir para a maneira como concebemos a habilidade humana de aprender e de perceber o mundo quando lidamos com sistemas artificiais que supostamente “aprendem” sobre nós. Isto é, procuro investigar possibilidades para a conservação da autonomia humana, no sentido

¹³⁰ Assumo que sistemas de machine learning não são cognoscentes; mais adiante explicarei por quê.

¹³¹ Rolla (2021) propõe a tradução “produção participativa de sentido”.

proposto pelo enativismo, num contexto em que a hipótese de autonomia para os sistemas costuma ser a que mais amplamente mobiliza a atenção¹³².

Minha preocupação principal é a de que esses sistemas, alimentando-se de dados gerados pelos seus próprios usuários, que depois geram mais dados e daí por diante, acabem restringindo seus utilizadores a uma rede fixa de interações e possibilidades, numa circularidade que pode ser danosa para a autonomia dos cognoscentes, reduzindo-a. Desse modo, se poderia constatar uma ameaça não apenas à autonomia em níveis mais básicos, por assim dizer, mas à capacidade de quem usa esses sistemas, aqui tomados como seres cognoscentes, de expandir seu conhecimento e suas experiências para além desses universos delimitados. Mais uma vez reitero que a educação, aqui, é concebida no sentido da formação humana, impactada de várias e significativas maneiras, hoje, pelos sistemas de aprendizagem de máquina. Acredito ainda que, ao investigar sobre a preservação da autonomia humana num mundo que tem valorizado as habilidades dos sistemas artificiais, bem como sua possível autonomia, e privilegiado as discussões nesse sentido, podemos recuperar alguns dos elementos essenciais a uma discussão sobre aprendizagem.

A perspectiva enativista contribui para esse resgate, como aqui espero demonstrar. Afinal, sob o entendimento do paradigma enativista para a cognição (Stewart, Gapenne & Di Paolo, 2010), concebe-se o aprendiz ou cognoscente como um agente capaz de fazer boas escolhas que, em última instância, garantem a sua integridade e identidade. O cognoscente deve construir sua própria perspectiva para o pensar, sendo um sujeito capaz de se desenvolver continuamente, transformando-se e sendo aquele que desenha o seu caminho (ainda que haja sempre limites e restrições inerentes ao meio, do mundo que já existe). Autonomia envolve, ainda, ser capaz de compartilhar com outros, já que o ambiente é constituído em parceria com outros, com quem as constantes trocas que fazemos são não uma opção, mas uma necessidade desde sempre enraizada na fonte das nossas capacidades cognitivas.

¹³² Neste ponto, vale comentar que é de se esperar que a possível autonomia das máquinas mobilize mais atenção do que possíveis ameaças à autonomia humana se pensarmos que vivemos em um mundo cérebro-centrado; como demonstrei no capítulo 3, pesquisas sobre o cérebro e para o cérebro prevalecem, encontram mais espaço (e recebem mais financiamento ou são financiadas com mais facilidade pelas agências), e tudo isso demonstra como talvez estejamos privilegiando a perspectiva do artificial em detrimento do humano.

Na seção a seguir irei introduzir o conceito de *participatory sense-making*, uma ideia chave para a argumentação presente neste capítulo.

5.1

O conceito de *participatory sense-making*

Para além de propor uma nova maneira de conceber a relação entre o sujeito cognoscente e o mundo, o enativismo tal como elaborado por De Jaegher e Di Paolo (2007) postula uma nova abordagem à cognição social¹³³. O conceito de *participatory sense-making* é o ponto de partida para essa concepção. A ideia é que, da interação entre (pelo menos) dois indivíduos, algo novo emerge, que não existiria se essa interação não acontecesse. A partir do momento em que há essa faísca, isto é, esse elemento único que surge somente porque há tal interação entre dois seres cognoscentes, os resultados dessa interação passam a definir o curso da própria interação – isto é, se ela segue, se é interrompida, como ficam a sua fluidez e o seu equilíbrio.

Participatory sense-making é, assim, a coordenação de atividades intencionais presentes em uma interação, em que processos individuais de *sense-making* são afetados, e novos domínios de *sense-making* social podem emergir que não estariam presentes em ações individuais (De Jaegher & Di Paolo, 2007, p. 497) – esta ideia segue o conceito enativista de emergência. Por conta disso é que o conceito de *participatory sense-making* estende o conceito de *sense-making*; este último é o acoplamento geral com o mundo protagonizado pelo agente, enquanto o *sense-making participativo* é o engajamento com outro(s) agente(s) para que se faça sentido do mundo de modo coletivo. Tecnicamente, interação, no enativismo, refere-se ao acoplamento entre o agente e um elemento particular do mundo: outro agente.

Uma abordagem enativa da compreensão social começa a partir do estudo da interação e da coordenação. A **interação** é aqui entendida como o acoplamento entre um agente e um aspecto específico de seu mundo: um outro agente. Interação é a interdependência mútua (ou o acoplamento bidirecional, corregulado) dos comportamentos de dois **agentes sociais**. (Di Paolo, Rohde & De Jaegher, 2010, p. 60, grifos meus)¹³⁴

¹³³ Ver De Jaegher e Di Paolo (2007) e De Jaegher e Di Paolo (2008).

¹³⁴ No original: *An enactive approach to social understanding starts from the study of interaction and coordination. Interaction is here understood as the coupling between an agent and a specific*

A interação passa a constituir um sistema com autonomia própria, autorregulada. Sendo assim, é importante destacar que, dentro desse sistema que deve manter a si mesmo, também a autonomia de cada agente deve ser mantida (ainda que possa ser reduzida) sob o risco de extinguir a interação caso isso não aconteça. Se deixa de haver autonomia de um dos lados, voltamos a ter *sense-making*, que é a interação com o mundo não social. Nesse caso, o “outro” equivaleria a “uma ferramenta, um objeto ou um problema para sua cognição individual” (De Jaegher & Di Paolo, 2007, p. 492).

A interação social é o acoplamento regulado entre pelo menos dois agentes autônomos, onde a regulação é direcionada a aspectos do acoplamento em si, de modo que constitua uma organização autônoma emergente no domínio da dinâmica relacional, sem destruir, no processo, a autonomia dos agentes envolvidos (embora o escopo deste último possa ser ampliado ou reduzido). (De Jaegher & Di Paolo, 2007, p. 493)¹³⁵.

Quando procuram definir interações sociais, De Jaegher e Di Paolo (2007); Di Paolo, Rohde e De Jaegher (2010) estão preocupados em oferecer um contraponto a abordagens tradicionais como a Teoria da Mente e a Teoria da Simulação. Essas abordagens concebem a mente como algo hermético que, por conta desse fechamento e opacidade, torna os indivíduos tais como quebra-cabeças para outros indivíduos. Isto é, para decifrarmos alguém, teríamos que fazer uma espécie de investigação científica (De Jaegher & Di Paolo, 2007). Segundo a linha de pensamento da Teoria da Mente, seria necessário aplicar inferências lógicas a grupos de conhecimentos e percepções; para as teorias defensoras da simulação, seria necessário simular o comportamento de um indivíduo para descobrir o que se passa com ele.

Como explica Colombetti (2014, p. 169), existem várias teorias que combinam essas ideias (e ainda aquelas advindas da TT – *Theory-Theory*, também baseada no pressuposto de que o entendimento do outro se dá por meio de um

aspect of its world: another agent. Interaction is the mutual interdependence (or bidirectional, co-regulated coupling) of the behaviors of two social agents. (Di Paolo, Rohde & De Jaegher, 2010, p. 60, grifos meus)

¹³⁵ No original: *Social interaction is the regulated coupling between at least two autonomous agents, where the regulation is aimed at aspects of the coupling itself so that it constitutes an emergent autonomous organization in the domain of relational dynamics, without destroying in the process the autonomy of the agents involved (though the latter's scope can be augmented or reduced).* (De Jaegher & Di Paolo, 2007, p. 493).

processo de inferência)¹³⁶. A premissa básica, e que aqui nos interessa para a elaboração do argumento proposto, é que os estados mentais do outro são ocultos, privados – por isso é que, para compreender o outro, é necessário alcançar esses estados via processos inferenciais. “Isso é também conhecido como a premissa de que entender outros é essencialmente uma questão de ler mentes”¹³⁷ (Colombetti, 2014, p. 171-172). Um ponto comum entre essas teorias, e que contrasta então com o enativismo, é que elas têm como pressuposto uma visão internalista, do indivíduo para dentro; trata-se da lógica cognitivista. Como explicitado em Di Paolo, Rohde e De Jaegher (2010, p. 60), sob essa perspectiva somos confrontados, numa situação social, com um outro que nos é impenetrável, inacessível e misterioso, nos restando para decifrá-lo recursos como o nosso raciocínio e a nossa imaginação. Além disso, o corpo não desempenha nenhum papel particularmente relevante nessas interações, e as questões acerca da autonomia, emergência, auto-organização etc., objetos de atenção do enativismo, não são tangenciadas. Há um privilégio da perspectiva do processamento de informações, de onde viria o sentido e o significado das coisas.

Mas, como esclarecem De Jaegher e Di Paolo (2007), essas teorias mais tradicionais têm enfrentado oposições, e tem crescido a ideia de que a racionalização sobre os problemas e a ênfase na capacidade de fazer previsões sobre o outro podem não estar no centro das atividades de caráter social. Compreender o outro seria uma forma de “prática corporificada” (Gallagher, 2001, p. 85, apud Colombetti, 2014, p. 172). Em vez de tentar adivinhar o que o outro sente, imaginar, fazer suposições, pode-se – adotando um olhar da fenomenologia para isto – conceber que simplesmente vivenciamos o outro por meio da percepção. “Percebemos a vida e a subjetividade no corpo do outro”¹³⁸ (Colombetti, 2014, p. 173). É daí que vem a conexão que é típica de um encontro entre dois agentes cognoscentes. Desse modo, o problema de descobrir o que se passa com o outro, ou perceber o que o outro sente, recebe outro tratamento na fenomenologia, alternativo à ideia de uma mente externa hermética e inacessível. Uma vez que o enativismo se apropria dessas noções, vale trazer aqui a concepção fenomenológica de empatia

¹³⁶ Não cabe aqui aprofundar as ideias dessas teorias, apenas citá-las para mostrar o quanto suas premissas se aproximam do cognitivismo e se afastam das concepções enativistas.

¹³⁷ No original: *This is also known as the assumption that understanding others is essentially a matter of mind reading or mentalizing.* (Colombetti, 2014, p. 171-172).

¹³⁸ No original: *we usually just perceive life and subjectivity in the other's body* (Colombetti, 2014, p. 173).

como parte do entendimento que se faz necessário para completar a explanação sobre intersubjetividade e *participatory sense-making*. Comumente, diz-se que empatia é a capacidade de “se colocar” no lugar do outro. Na perspectiva da fenomenologia, empatia é um “acesso experiencial à subjetividade do outro”, como coloca Colombetti (2014, p. 174) – a autora recorda que o termo inglês “*empathy*” significa “*feeling in*”.

Vamos imaginar que uma pessoa (A) esteja gritando, e outra pessoa (B) infira que o motivo pelo qual a outra está gritando é porque sente dor. Isso se dá porque, acostumada a gritar quando sente dor, a pessoa (B) infere que (A) esteja sentindo algo semelhante quando reage assim. A pessoa (B) pressupõe, ainda, que a sua própria mente é a única a que ela em acesso direto, e que aquele corpo exterior a (B) está, provavelmente, ligado de maneira semelhante a uma mente exterior, a mente de (A). Portanto, (B) não teria acesso à mente de (A). Este exemplo, em Zahavi (2021, p. 126), descreve a maneira clássica de lidar com o problema da intersubjetividade, conhecida, como o autor indica, como *argumento da analogia*. No parágrafo a seguir, Zahavi coloca duas questões que se opõem à abordagem da intersubjetividade como uma questão de analogia e inferência:

Primeiro de tudo, se poderia questionar a alegação de que minha experiência de mim mesmo é de uma natureza puramente mental, ocorrendo de forma isolada da experiência de outros, e a precedendo. Em segundo lugar, o argumento da analogia assume que nós nunca experimentamos os pensamentos ou os sentimentos de outra pessoa, mas que somente podemos inferir sua provável existência com base naquilo que nos é dado, nomeadamente um corpo físico. Mas, por um lado, essa suposição parece implicar em uma abordagem muito intelectualista – afinal, animais e crianças parecem compartilhar a crença em outras mentes, porém, no caso deles, dificilmente isso é resultado de um processo de inferência – e, por outro lado, parece pressupor uma dicotomia altamente problemática entre o interior e o exterior, entre experiência e comportamento. (Zahavi, 2021, p. 127)

Uma proposta alternativa, fenomenológica, é olhar para o problema a partir da experiência que se tem do outro. Nós vemos a emoção no rosto do outro:

[U]ma solução para o problema das outras mentes deve começar com um entendimento correto da relação entre corpo e mente. Em algum sentido, experiências não são internas, não se escondem na cabeça, mas estão expressas em gestos corporais e ações. Quando vejo a face de alguém, eu a vejo como amigável ou zangada etc.; isto é, a própria face expressa essas emoções. Além disso, o comportamento corporal é carregado de significado, é intencional e, dessa forma, não é interno ou externo, mas está para além dessa distinção artificial. (Zahavi, 2021, p. 127)

Desse modo, quando olhamos para alguém, nós não primeiro percebemos um corpo físico para, num movimento posterior, depreender a subjetividade alheia; não inferimos a existência do outro. “No encontro face a face, nós não somos confrontados nem com um mero corpo, nem com uma psique escondida, mas com um todo, unificado. Nós vemos a raiva do outro, sentimos sua tristeza” (Zahavi, 2021, p. 127). No lugar de uma inferência por analogia existe, segundo o autor, um modo distinto de consciência que se chama empatia ou *Fremderfahrung*. É a empatia, nessa apreciação do conceito, que nos permite “experimentar os sentimentos, desejos e crenças dos outros, de uma maneira mais ou menos direta” (Zahavi, 2021, p. 127). A habilidade de interagir com o outro, corporificado, e de reconhecê-lo como uma subjetividade externa é algo que se torna possível justamente por conta do fato de que a nossa subjetividade é corporificada. Esse é o recurso que nos permite antecipar a subjetividade do outro (Zahavi, 2021, p. 129-130). “Nossas experiências de outras pessoas, mesmo quando as observamos remotamente, são fundamentadas na intersubjetividade corporificada, e são, assim, um aspecto inerentemente *social* da experiência pessoal, oposto à racionalização que internamente suplementa a imagem do outro com inferências sobre sua presumida mentalidade”¹³⁹ (Di Paolo, Cuffari & De Jaegher, 2018, p. 63).

O primeiro corpo que experimentamos é o nosso próprio; quando uma de nossas mãos toca a outra, por exemplo, isso nos permite perceber o que Zahavi chama de *natureza dual do corpo*, sendo a interioridade e a exterioridade manifestações de um mesmo fenômeno.

Minha autoexploração corporal me permite confrontar com *minha própria* exterioridade, e o argumento tem sido de que essa experiência é uma precondition crucial para a empatia. Sou capaz de encontrar um outro porque a minha experiência de mim mesmo já contém, em si, uma dimensão da alteridade (Zahavi, 2021, p.130).

É interessante notar que, apesar de haver um acesso direto ao outro proporcionado pelo próprio fato de a subjetividade corporificada o permitir, ainda é mantida nesse processo uma certa inacessibilidade a ele. Isso é inerente ao próprio fato de o outro ser o outro, isto é, ter a sua individualidade; a alteridade do outro não pode se dissolver exatamente porque isso extinguiria o problema da

¹³⁹ No original: *Our experiences of other people, even when we observe them remotely, are grounded on embodied intersubjectivity, and thus are an inherently social aspect of personal experience, as opposed to a rationalization that internally supplements the image of the other with inferences about their presumed mindedness* (Di Paolo, Cuffari & De Jaegher, 2018, p. 63).

intersubjetividade. Por outro lado, “a discussão fenomenológica da subjetividade, que é uma discussão do sujeito transcendental ou constitutivamente em funcionamento, se torna uma discussão não simplesmente do Eu, mas do *nós*”. (Zahavi, 2021, p.138).

Para propor uma nova maneira de abordar a interação social é que De Jaegher e Di Paolo (2007) concebem o conceito de *participatory sense-making*. Refletindo a influência da fenomenologia sobre a teoria enativista, o conceito admite a importância da expressividade no contato social, mas não se restringe a isso: há uma ênfase no processo de interação em si, que conta com uma dimensão afetiva, em que o grau de conexão com o outro pode ser sentido em vários níveis. O foco está no acoplamento entre dois agentes, tendo como base a reciprocidade corporal e coordenada. É proposta uma concepção de intersubjetividade calcada em capacidades primárias, que não demandam conceitos, como a capacidade de imitação, de distinguir entre objetos inanimados e pessoas e ainda a capacidade de ser responsivo às expressões faciais do outro, como coloca Colombetti (2014, p. 172).

Quando ocorre *participatory sense-making*, o significado ou o sentido das coisas é gerado e transformado pelos agentes enquanto eles interagem, não é *transportado* entre eles como se fosse um objeto ou algo pronto e acabado – como ressalta De Jaegher (2009). Citando Tronik et al. (1979), De Jaegher e Di Paolo (2007) mencionam o exemplo de como um bebê se mostra incomodado quando sua mãe, que sempre interage bastante com ele, subitamente se senta imóvel, com uma face neutra – esse caso poderia ser explicado por meio da expressividade, que de repente foi interrompida; mas em outro caso, semelhante, se percebe a falta não de expressividade, mas de uma interação que seja contínua, o que fortalece o argumento da conexão típica do processo de interação em si. Esse outro caso (em Murray & Trevarthen, 1985, apud De Jaegher & Di Paolo, 2007) é o do bebê que se sente perturbado quando uma transmissão ao vivo de sua mãe, numa tela, é substituída por uma gravação dela na mesma interação, porém momentos antes. Isto é, há a expressão da mãe na tela, mas não é o que está acontecendo naquele momento mais; assim, o bebê se incomoda porque sente a perda da conexão que somente poderia ser estabelecida no processo corrente de interação entre os dois.

Para além do que já explicitiei sobre essa ideia de emergência compartilhada de sentido, vale ressaltar sua proximidade com as dinâmicas de coordenação

presentes nas teorias de sistemas dinâmicos. A lógica desses estudos permite uma análise estendida no tempo que, aplicada às interações sociais, pode ser útil para que sejam observadas as oscilações pertinentes a essas reorganizações e a esses reequilíbrios constantes.

[P]assamos algum tempo descrevendo o fenômeno da coordenação entre sistemas acoplados. Isto nos permite conceber as interações como processos estendidos no tempo com uma estrutura rica que só é aparente no nível relacional da dinâmica coletiva. Quando compreendermos como a coordenação surge, é sustentada, muda e se rompe durante os encontros sociais, estaremos em condições de estabelecer uma conexão entre estes aspectos temporais da interação. (De Jaegher & Di Paolo, 2007, p. 490)¹⁴⁰

Ainda segundo De Jaegher e Di Paolo (2007, p. 490), a coordenação é definida nesse contexto como a correlação não acidental entre os comportamentos de dois ou mais sistemas em acoplamento sustentado, ou que estiveram acoplados entre si no passado ou a outro sistema, comum. Correlação é a coerência observada entre dois ou mais sistemas acima daquilo que é esperado, considerando aquilo que os sistemas são capazes de fazer. Os autores dão o exemplo de uma multidão de pessoas que eventualmente caminha para o mesmo lado na rua sem “combinar” aquilo previamente; poderia haver um fator de coordenação disso, como por exemplo se elas estivessem em uma manifestação por uma determinada causa comum¹⁴¹.

Qual a importância de olhar para o conceito de coordenação se queremos compreender as interações sociais? É essa noção que, na perspectiva do enativismo, explica por que o encontro é que é capaz de sustentar-se a si mesmo (ao próprio encontro); isto é, “padrões de coordenação podem influenciar diretamente a disposição contínua de indivíduos envolvidos em sustentar ou em modificar seu encontro” (De Jaegher & Di Paolo, 2007, p. 491). Presenciamos mútuas regulações a todo o tempo em nossas interações com as pessoas, sejam elas curtas ou longas, durante as quais percebemos que nossos atos e os atos delas ora se acomodam ora

¹⁴⁰ No original: *[W]e spend some time describing the phenomenon of coordination between coupled systems. This allows us to view interactions as processes extended in time with a rich structure that is only apparent at the relational level of collective dynamics. Once we understand how coordination arises, is sustained, changes, and breaks down during social encounters, we will be in a position to make a connection between these temporal aspects of interaction. (De Jaegher & Di Paolo, 2007, p. 490)*

¹⁴¹ Outros exemplos de correlação, em sistemas animais, físicos e biológicos, são dados em De Jaegher e Di Paolo (2007, p. 490).

entram em atrito, definindo os próximos passos, que vão definir os demais e daí por diante. São diversas camadas de emergência.

Casos bastante cotidianos servem para ilustrar essas mútuas regulações sociais, como uma conversa com um amigo em que percebemos, por seus gestos, tom de voz ou expressões faciais, que está incomodado por algum motivo, e assim podemos perguntar o que há ou tentar mudar o curso da interação de outra maneira; numa sala de aula, um professor pode, a partir da observação do comportamento de sua turma, reorganizar o que havia planejado para aquele dia, se notar que isso se faz necessário; uma mãe vai descobrindo a melhor maneira de amamentar seu bebê na medida em que experimenta essa ação e explora a conexão junto a ele e daí por diante.

Todas as situações de interação social como essas têm em comum o fato de que são os próprios elementos que emergem delas que definem a maneira como elas vão (ou não vão) continuar se desdobrando; os elementos que são gerados nesses contextos são os únicos capazes de sustentar (ou não) as situações. Esses aspectos não estão em um *ou* em outro agente social que participa dos processos, mas justamente nas engrenagens entre eles. Se as interações são interrompidas em algum momento, isso revela que os elementos presentes nessas engrenagens não puderam seguir coordenados de uma forma capaz de sustentá-las. Num certo sentido, a mesma lógica vale para o acoplamento individual com o ambiente: ele deve continuar para que suporte a si mesmo, enraizando-se a nas articulações e trocas entre aquilo que é interno ao agente e aquilo que é externo a ele. Então, a capacidade de sustentar esse acoplamento não está unicamente no ser cognoscente *ou* no seu ambiente, mas nesse intercâmbio contínuo que permite a manutenção da vida do ser vivo. A habilidade de interagir emerge da própria interação. Não é possível conceber de maneira prévia os indivíduos como “interatores”; “indivíduos coemergem como interatores na interação”¹⁴² (De Jaegher & Di Paolo, 2007, p. 492). De maneira complementar, se a autonomia de cada parte precisa necessariamente ser mantida durante toda a interação, ou irá rompê-la, parece que uma dança é um caso eficaz para a compreensão definitiva desse preceito, como De Jaegher e Di Paolo (2007) sugerem:

¹⁴² No original: *individuals co-emerge as interactors with the interaction*. (De Jaegher & Di Paolo, 2007, p. 492).

Uma dança em casal envolve mover um ao outro, fazer o outro se mover e ser movido pelo outro. Isto vale tanto para o líder quanto para o seguidor. Seguir (um ao outro) faz parte de um acordo e não equivale a ser mudado de posição pelo outro. Se quem segue desistisse de sua autonomia, a dança em casal terminaria ali, e a situação se pareceria mais com uma boneca sendo carregada pela pista de dança. O mesmo vale para as conversas: cada parceiro deve se engajar a partir de um ponto de vista autônomo (De Jaegher & Di Paolo, 2007, p. 494)¹⁴³.

O conceito de *participatory sense-making* é carregado do sentido de afetividade que permeia a concepção enativista para a cognição. Para além de se referir às emoções sentidas de afeto, do gostar ou ter estima pelo outro, afetividade no enativismo tem uma outra concepção. Conforme explicitado no capítulo quatro, dedicado ao enativismo e seus conceitos, afetar-se no sentido enativista significa ser afetado(a) pelo mundo e pelos outros; é a noção de que nada que é constitutivo do ambiente nos é indiferente. Essa ideia pode ser relacionada com a concepção fenomenológica de empatia, já mencionada, que está ligada à noção de intersubjetividade corporificada. Interessante notar que, uma vez que contamos com algo que a fenomenologia chama de *tendência empática* (*empathic tendency*; Colombetti, 2014, p. 178) – a “tendência ou a tentação de ver sentimentos em qualquer coisa que reproduza características similares a aquelas que em humanos expressam experiências reais” – nós podemos sentir esse impulso com relação a algo que não é humano e com o qual não temos uma conexão de fato. A tendência empática pode ser observada, por exemplo, com relação a robôs. Aqueles que reproduzem características humanas visuais e vocais se tornam “irresistíveis” a essa tendência. Segundo a autora indica, isso foi observado no robô Kismet¹⁴⁴, desenvolvido por Cynthia Breazeal no MIT (Breazeal, 2002, apud Colombetti, 2014). Kismet tem uma face grande e um corpo pequeno; é capaz de mexer o pescoço, os olhos, as sobrancelhas, os lábios – que tomam formas variadas, entre outros movimentos. Colombetti destaca que, por mais que os usuários que olham para o robô saibam que ele não está sentindo ou experimentando nada, ainda assim não podem controlar a sua tendência empática, geralmente demonstrando compaixão quando o robô exibe uma cara triste ou tristeza quando ele é desligado.

¹⁴³ No original: [C]ouple dancing involves moving each other, making each other move, and being moved by each other. This goes for both leader and follower. Following is part of an agreement and does not equate with being shifted into position by the other. If the follower were to give up her autonomy, the couple dancing would end there, and it would look more like a doll being carried around the dance floor. The same goes for conversations: each partner must engage from an autonomous standpoint (De Jaegher & Di Paolo, 2007, p. 494).

¹⁴⁴ Ver <https://www.youtube.com/watch?v=Kw-gOmJwzuc>

Mesmo que o robô permanecesse ligado, no entanto, não haveria uma real conexão a ser sustentada, já que o outro lado, do sistema artificial, não tem essa tendência empática. Ideias como esta receberão mais atenção nos tópicos seguintes.

5.2

Participatory sense-making e machine learning

Analisar a perspectiva do enativismo para a interação entre dois indivíduos, ou agentes, pode ser uma boa chave para se compreender a relação humana com sistemas de machine learning. Penso que o exemplo da dança entre um casal, mencionado anteriormente, sugere o raciocínio a seguir: quando se engajam em uma interação como tal, parceiros de dança são agentes em equilíbrio, coordenados, já que sua autonomia é sustentada durante o processo. Isso se mantém ainda que um seja um dançarino mais experiente e o outro, menos hábil, precise seguir seus passos e movimentos em alguma medida. Em linha com esse raciocínio, se poderia perguntar: qual é a distribuição do “peso” desse equilíbrio quando o “parceiro de dança” de um ser humano é um sistema artificial de machine learning? Somos como marionetes “levadas” por esses sistemas? Esses sistemas é que seriam “levados” por nós? Talvez nenhuma dessas alternativas; mas, ainda assim, valeria questionar: podemos, então, fazer movimentos diferentes dependendo de como esse acoplamento acontece, ou de quanto nós sabemos sobre esse acoplamento ou sobre o “parceiro”, de modo, então, a gerenciarmos isso para podermos cooperar para ajustar o processo? E esse processo, de alguma maneira, afetaria a nossa autonomia enquanto *sense-makers*?

Na esteira desse questionamento vêm outros, um dos quais é recorrente e já foi mencionado: os sistemas de machine learning podem ser considerados autônomos? Ainda que essa questão levante ainda outras, possivelmente para além do escopo desta tese, demandando em si extenso escrutínio, proponho que antes de prosseguir a analisemos brevemente sob lentes enativistas. Segundo Thompson (2007), p. 44), *autonomia* pode ser observada em um sistema sem a organização bioquímica da célula, e sem uma membrana feita de matéria. É o caso de uma colônia de insetos, que formam uma rede cuja fronteira ou “membrana” é social e territorial, não material como a da célula. *Autonomia* também poderia ser observada em um sistema artificial, segundo Thompson (2007, p. 50). Mas um aspecto é

primordial aqui: seres ou sistemas autônomos se distinguem daqueles sistemas que são definidos a partir de seu lado exterior, que são os sistemas heterônomos. A citação a seguir, de Rohde (2010), ajuda a deixar esse argumento mais claro:

As restrições impostas aos processos de geração de identidade auto mantidos são de natureza mecânica. Os organismos vivos estão vinculados pelas leis da física, mas as possibilidades de se reorganizarem e, com eles, o mundo de interações significativas que eles geram, são abertas. Esta abertura contrasta com o desenho explícito de circuitos adaptativos nas abordagens computacionalistas, por exemplo, na disciplina de aprendizagem de máquinas. (Rohde, 2010, p. 20)¹⁴⁵

Sistemas de machine learning são desenvolvidos a partir de certos parâmetros, os quais se encontram codificados nos algoritmos. Concebe-se que, na medida em que rodam, esses sistemas “aprendem”, o que significa que dados internos a eles se tornam material bruto para gerar mais dados (sem dados, esses tipos de sistemas não se mantêm). Ainda que, inicialmente, sejam alimentados por dados que vêm de fora deles, isto é, por atividade humana externa aos sistemas, dados também emergem da operação interna dos sistemas, na medida em que eles são utilizados, o que configura um movimento que conta com um aspecto circular.

Isso posto, o problema é: esses sistemas seriam considerados heterônomos ou autônomos? É algo que tende a depender da perspectiva adotada. O paradigma para definir as trocas entre um sistema heterônomo e o meio que o circunda é o paradigma de *inputs* que vêm de fora para serem processados internamente ao sistema para gerar *outputs*. Quando o resultado ou *output* esperado não acontece, compreende-se que houve um erro no sistema. Já o paradigma para um sistema autônomo, por outro lado, é o de uma conversa, na qual resultados não satisfatórios são vistos como quebras no entendimento (Thompson, 2007).

Em princípio, um sistema conexionista tenderia a ser considerado heterônomo; no entanto, as redes neurais artificiais que constituem os sistemas de machine learning não são somente alimentadas por informação externa. Por isso, parece que sistemas de machine learning oscilam entre esses dois extremos. No entanto, como Rohde (2010, p. 20) explica, “Mesmo que machine learning seja um campo que está florescendo na engenharia de software, algoritmos estão sujeitos à

¹⁴⁵ No original: The constraints imposed on self-maintaining processes of identity generation are of a *mechanical* nature. Living organisms are bound by the laws of physics but the possibilities to reorganise themselves and, with them, the world of meaningful interactions they bring forth, are open-ended. This open-endedness contrasts with the explicit design of adaptive circuits in *computationalist approaches*, e.g., in the discipline of machine learning. (Rohde, 2010, p. 20).

limitação de regras internas ao sistema”¹⁴⁶. Isto é, as regras existem nos sistemas antes de eles começarem a operar; foram previamente estabelecidas; sendo assim, é improvável que novas regras sejam geradas durante a operação desses circuitos. Por outro lado, no curso das interações humanas, novas regras ou normas são criadas e antigas são reavaliadas ou postas em xeque a todo momento.

Outro aspecto a ser ressaltado da discussão acerca de como abordar a relação humana com sistemas de machine learning (no sentido de procurar compreender se pode ser considerada *participatory sense-making*) é o da *corporificação*. Como Thompson (2007) coloca, a mente humana emerge de processos de auto-organização que interconectam intimamente o cérebro, o corpo e o ambiente em vários níveis. Como Di Paolo, Rohde e De Jaegher (2010) propõem, o corpo a partir do qual emerge a mente é, em si, um sistema animado, com muitas camadas autônomas de autoconstituição, autocoordenação e auto-organização em vários níveis de abertura ao mundo. Seguindo esse raciocínio, sistemas de machine learning poderiam ter uma mente como a humana somente se contassem com corpos imersos no ambiente e com uma organização tipicamente encontrada em seres humanos¹⁴⁷. Até o momento, somente criaturas corporificadas, orgânicas, com capacidades sensório-motoras e habilidades intersubjetivas poderiam ser consideradas agentes cognitivos sob essa perspectiva¹⁴⁸. Na falta de um corpo com as dimensões orgânica, sensório-motora e intersubjetiva, características

¹⁴⁶ No original: *Even if machine learning is a blossoming field as part of software engineering, such algorithms are functionally constrained by in-built rules.* (Rohde, 2010, p. 20)

¹⁴⁷ Uma das características mais importantes para o enativismo é a organização do ser e suas estratégias para sustentá-la enquanto está enraizado no ambiente; assim, se fosse possível instanciar a organização de um ser vivo/a cognoscente/criador de sentidos em um sistema artificial, seria possível ter uma mente artificial. Mas ao ser ainda faltaria uma história vivida, uma ideia que está ligada ao conceito enativista de becoming; e isso tornaria mais difícil para um agente artificial tornar-se uma *pessoa*. (ver Di Paolo, 2020).

¹⁴⁸ Di Paolo e Thompson (2014) dizem que "Na medida em que os robôs não conseguem exibir a dinâmica dos sistemas vivos, eles não são agentes autônomos, autorregulados, e seu sense-making não está fisicamente fundamentado em processos autopoieticos e metabólicos. Portanto, há boas razões para pensar que eles não podem ter emoções como as nossas e são incapazes de dar sentido ao seu ambiente através de um enquadramento afetivo. Isto inevitavelmente torna suas formas de interpretar seu ambiente muito diferentes do sense-making que criaturas como nós realizam regularmente". Zebrowski e McGraw (2021), contudo, discutem uma abordagem num outro sentido, de artefatos interpretados como robôs que, baseados em hábitos, poderiam se tornar sistemas genuínos de criação de sentido.

*No original: *that To the extent that robots fail to exhibit the dynamics of living systems, they are not self-regulating, autonomous agents, and their sense-making is not physically grounded in autopoietic and metabolic processes. Thus, there is good reason to think that they cannot have emotions like ours and are incapable of making sense of their surroundings via affective framing. This inevitably makes their ways of interpreting their surroundings very different from the sense-making that creatures like us carry out on a regular basis.*

mandatórias para terem cognição no sentido enativista, aos sistemas artificiais de machine learning faltaria ainda um propósito intrínseco comumente encontrado em sistemas autônomos (Colombetti, 2014). Sendo assim, talvez sistemas de machine learning possam apresentar algum nível de autonomia, já que, ainda que não tenham um corpo tal como o enativismo define corpo, são capazes de sustentar um equilíbrio entre seu interior e exterior que os mantêm em operação. No entanto, tende a não ser possível dizer que seria o mesmo se relacionar, ou “dançar”, com um sistema de aprendizagem de máquina ou com outro ser humano. Esta argumentação será complementada na seção seguinte.

5.3

Ao lidar com sistemas de machine learning, seguimos sendo *sense-makers* autônomos?

Por todas as razões que expus até o momento, parece adequado dizer que pode haver algum nível de *autonomia* num sistema de machine learning. No entanto, parece também coerente dizer que não há *sense-making participativo* quando nós, humanos, sempre prontos a nos tornar *interatores*, nos engajamos com esses sistemas. Falta-lhes as condições para serem considerados agentes, *sense-makers* e, consequentemente, *interatores*. Isso se dá especialmente porque a atividade de fazer sentido do mundo recruta a autonomia, a adaptatividade e o corpo como um todo, bem como a afetividade, elementos que ainda se mostram presentes quando humanos lidam com sistemas de machine learning, mas se encontram, claro, apenas nos humanos, não nos sistemas artificiais. *Sense-making participativo* exige pelo menos dois corpos, duas mentes; dois agentes ou *sense-makers* que, no encontro, se tornam *interatores*. Em resumo, demanda ao menos dois seres autônomos ou *sense-makers*. Como De Jaegher e Di Paolo (2007, p. 498) esclarecem: “[e]ncontramos situações em que, através da coordenação do sense-making, um dos interlocutores é orientado para um novo domínio de significado que fazia parte da atividade de sense-making do outro”¹⁴⁹. Essa citação nos revela como o *sense-making participativo* leva ao inesperado, abre espaço para a surpresa aparecer, de uma maneira que somente pode ocorrer quando dois *sense-makers* individuais e

¹⁴⁹ No original: [W]e encounter situations where, through coordination of sense-making, one of the interactors is oriented towards a novel domain of significance that was part of the sense-making activity of the other. (De Jaegher & Di Paolo, 2007, p. 498)

autônomos interagem. Além dessa surpresa presente na emergência de algo novo que somente passa a existir por conta dessa interação, há o fator do imprevisível inerente ao sistema cognitivo: o imprevisível conecta-se antes de tudo ao fato de que a individuação é um processo precário e contínuo (Di Paolo, 2009).

No intuito de me ater à questão proposta no título desta seção, questiono: pode-se esperar que o mesmo nível de *imprevisibilidade*, por assim dizer, seja encontrado quando lidamos com sistemas de machine learning? Provavelmente não. Mas como isso impacta nossa *autonomia*? Se a possibilidade do inesperado não der conta de responder sozinha pela presença ou não de autonomia, pode ao menos responder por parte disso. Afinal, o fator previsibilidade ou imprevisibilidade parece ser uma das chaves a guiar uma interação. Se ela for totalmente previsível, não é uma interação de fato, já que o que advém da interação é exatamente aquilo que vai defini-la. Se por um lado esse fator orienta o *sense-making participativo*, por outro ele é, em primeiro lugar, uma característica do *sense-making*.

Apresentando as diferenças entre o conceito de comportamento como um programa motor e como uma cinética em constante desenvolvimento, Sheets-Johnstone menciona uma caça como um bom exemplo para analisar a forma como nos movemos no mundo. Eu a menciono aqui por estar conectada a essa questão de previsibilidade:

Uma caça não é e não pode ser uma sequência específica e replicável de ações, tanto porque o mundo não é o mesmo de um dia para o outro ou possivelmente mesmo de momento para momento em relação ao terreno ou ao clima, por exemplo, como porque o movimento dos seres vivos não é o mesmo de um dia para o outro ou mesmo de momento para momento. De maneira essencial, é imprevisível. (Sheets-Johnstone, 2010, p. 176)¹⁵⁰.

Como numa eterna caça, ao agirmos no mundo não é possível prevermos por completo o que acontece em torno de nós, já que o contexto muda e também nós mudamos. De maneira semelhante, não parece ser possível prever tudo aquilo que irá emergir do encontro de dois seres humanos, já que o encontro em si vai fazer brotar algo que não despontaria a não ser graças à interação. Isso se estende à

¹⁵⁰ No original: *A hunt is not and cannot be a specific and repeatable sequence of actions, both because the world is not the same from one day to the next or possibly even from moment to moment with respect to terrain or weather, for example, and because the movement of living creatures is not the same from one day to the next or even from moment to moment. In essential ways, it is unpredictable. (Sheets-Johnstone, 2010, p. 176)*

linguagem, tratada no enativismo de uma maneira particular e, naturalmente, em linha com a lógica dinâmica da abordagem:

A linguagem viva é dirigida, significada e ressignificada, vitalizada, enquadrada e esculpida pela carnalidade dos engajamentos interativos entre pessoas reais de maneiras que não podem ser totalmente determinadas nem totalmente previstas pelas capacidades, intenções e experiências que trazem para o encontro, nem pelas limitações mais amplas das normas sociais, pela situação ambiental ou pelas regras gramaticais. O locus onde a linguagem é decretada é o meio termo frequentemente negligenciado entre a mente individual e os padrões socioculturais mais amplos (Di Paolo, Cuffari & De Jaegher, 2018, p. 132)¹⁵¹.

Vou a partir de agora seguir no desenvolvimento de meu raciocínio mostrando como, ao lidar com sistemas de machine learning, os cognoscentes se encontram em situações que não podem ser descritas como *sense-making participativo* e que, conseqüentemente,

- I. Na ausência de interação/de outro interator ou agente, por vezes não permitem ajustes mútuos entre os *interatores*, o que gera um encontro desbalanceado/desequilibrado que ameaça a autonomia plena do cognoscente;
- II. Tendem a reduzir os tipos de experiência vividos pelo cognoscente, conseqüentemente provocando uma queda nas chances de que ele se mantenha autônomo no mais amplo sentido possível de sua vivência/experiência no mundo;
- III. Configurando situações em que as normas/regras não são claras para o cognoscente, geram um cenário difícil para o cognoscente manter a sua autonomia, já que, para agir no mundo de uma maneira que a garantisse, ele precisaria estar ciente de regras que o orientem e potenciais riscos a evitar.

I. O encontro desequilibrado que ameaça a autonomia do cognoscente

¹⁵¹ No original: *Living language is directed, signified and resignified, vitalized, framed, and sculpted by the carnality of interactive engagements between real people in ways that cannot be fully determined nor fully predicted by the capabilities, intentions, and experiences they bring to the encounter, nor by the broader constraints of social norms, environmental situation or grammar rules. The locus where language is enacted is the often neglected middle term between the individual mind and the broader sociocultural patterns* (Di Paolo, Cuffari & De Jaegher, 2018, p. 132)

A interação social envolve os objetivos de (ao menos) dois indivíduos ou agentes – para usar uma linguagem enativista, em vez de objetivos poderíamos chamar de propósito (*purpose*). Ambos são orientados em direção a esses objetivos, de modo que o encontro, em si, envolve expectativas e planos de ambos os lados. Essas expectativas podem mudar durante o encontro, ou até se ajustar mutuamente durante o processo – como ressaltado, por exemplo, em De Jaegher e Di Paolo (2007) e De Jaegher (2009). Como consequência, o nível de autonomia que se consegue manter depende dos objetivos e do balanço observado na interação em que se pretende alcançá-los.

Vou estender meu argumento utilizando para isso mais um exemplo presente em De Jaegher e Di Paolo (2007, p. 500-501). Na situação que eles mencionam, que chamo aqui de situação (1), um jogo de mímica está transcorrendo. Uma pessoa, pertencente a um dos grupos de participantes, precisa fazer gestos representando algo que o grupo deve adivinhar o que é (por exemplo, o título de um filme). O ponto, portanto, é tornar fácil para o grupo descobrir o que se está gesticulando; afinal, assim o grupo marca pontos na competição. A pessoa responsável por fazer a mímica se esforça ao máximo para “aprender” a partir daquilo que o grupo está errando ao dar seus palpites; ela tenta melhorar seus movimentos, isto é, aprimorar a qualidade de sua mímica. O que acontece neste contexto? Enquanto essa pessoa ajusta sua atividade com base naquilo que os membros de seu grupo demonstram, procurando ajudar o grupo na adivinhação, os membros “adivinhadores” também vão modulando seus palpites a partir da evolução dos gestos de quem faz a mímica. Nessa situação, todos entendem que são um time e que seu objetivo comum é marcar pontos a favor desse time.

Agora vou explicar o que chamo de situação (2), que não é mencionada por De Jaegher e Di Paolo, mas que proponho aqui como uma extensão do exemplo dado por eles. E se, por alguma razão, o participante responsável por fazer a mímica desejasse confundir os membros de seu próprio grupo, em vez de levá-los a adivinhar a resposta correta? Ele provavelmente gesticularia de maneira diferente, trabalhando para que o grupo não acertasse e não marcasse pontos a favor da equipe.

Em ambos os exemplos, há um objetivo a ser alcançado e intenções envolvidas. O ponto onde gostaria de chegar é que sistemas de machine learning também têm seus “objetivos” atribuídos. Então, dentro do contexto do exercício que aqui proponho, cabe talvez perguntar: os objetivos desses sistemas parecem

operar mais como na situação (1) ou na situação (2)? Isto é, esses sistemas têm objetivos que se alinham com os de seus usuários, de modo a formarem um time que joga junto, ou tendem a ser mais como na situação 2, em que acaba por confundi-los?

Independentemente da resposta, considero pelo menos dois elementos importantes conectados a essas questões e que poderiam ser levantados nesta altura:

1) Apesar de serem criados para sustentar uma certa lógica que deve levar ao alcance de determinados objetivos, sistemas de aprendizagem de máquina não experimentam e nem sentem nada e não são, portanto, afetados pelo ambiente; não são *sense-makers*; a maneira como se “comportam” depende de como o sistema é programado e, portanto, das intenções e objetivos que estão por trás da programação. Há regras inseridas nos algoritmos;

2) Como o objetivo desta análise é considerar o que é que acontece com a autonomia do ser cognoscente que lida com sistemas de ML, poder-se-ia colocar que, num certo sentido, não haveria diferença entre um agente (humano) lidar com um sistema de ML ou lidar com um outro ser humano, sob o ponto de vista dos objetivos ou intenções. Afinal, ele poderia ajustar seus movimentos sendo o outro uma máquina artificial ou um ser humano como ele.

Mas, se olharmos cuidadosamente e considerarmos que o sistema pode *parecer* um interator/agente mas na verdade não é, devemos constatar que isso atrapalha o *sense-maker* em sua capacidade de se ajustar, também – no mínimo, a limita. Pois o *sense-maker* lida com um sistema que não é capaz de (e nem se preocupa em) fazer ajustes com base em percepções de limites, emoções, gestos, expressões faciais. Não há, portanto, uma situação de acomodações mútuas, de ajustes *como se estivesse em uma interação*, algo que demanda ao menos dois seres cognoscentes em conexão, de acordo com a proposta enativista de *sense-making participativo*. Esses sistemas seguem seus objetivos cegamente, indiferentes ao sentido ou ao *sense-maker* que podem estar afetando.

Sendo cegos ao agente, os sistemas de ML podem, por exemplo, confundir o cognoscente ou alcançar objetivos que dariam conta dos seus objetivos e necessidades, mas não dariam conta dos objetivos e necessidades (e nem dos sentimentos) do *sense-maker*. Não há como o *sense-maker* regular ou moldar seus movimentos se o outro lado não faz o mesmo; se não há conexão, fluxo de movimentos ou interação em curso, a situação se torna desbalanceada sob o ponto

de vista da autonomia enativista e o *sense-maker* pode ter dificuldade de se flexibilizar, se adaptar para encontrar o que seria melhor para a sua experiência – mesmo que isso ocorra inicialmente num nível subconsciente. Isso então ameaça a sua autonomia, ainda talvez não a anule por completo.

Em uma troca com um parceiro que não é, digamos assim, um seguro jogador do mesmo time, mas alguém em última instância cego para qualquer necessidade que possa emergir durante essa relação, o *sense-maker* encontra um ambiente pouco fluido para sua autonomia e adaptatividade. Um ambiente árido e inflexível, por ser desprovido de conexão e afetividade. Até aí, ele poderia contornar a situação, já que estamos sempre, como *sense-makers*, imersos em meios (sejam eles físicos, políticos, sociais, culturais) que podem nos ser inóspitos e sempre apresentam perturbações às quais precisamos nos ajustar para garantir nossa auto-organização e manutenção. No entanto, tendo em mente que os sistemas de ML são pensados nesta tese sob a lógica de possivelmente contribuírem para o processo de aprendizagem, vale pensar se seriam adequados dentro desse contexto e como e quando se poderia perspectivar seus usos com fins educacionais de modo a ampliar as possibilidades de autonomia, em vez de reduzi-la. Afinal, as relações e interações deveriam potencializar a cognição e ampliar as situações de aprendizagem, no lugar de reduzi-las. No ponto seguinte vou alargar um pouco mais o argumento.

II. Uma gama menor de experiências

A situação de lidar com sistemas de ML impacta de uma outra maneira a autonomia do cognoscente, para além da ameaça a partir do desequilíbrio apresentado no ponto anterior. Esta outra maneira nos leva de volta ao problema da previsão ou da surpresa, que abordei no começo desta seção. Além de dificilmente poderem ser considerados “neutros” – já que são criados e alimentados por humanos, e humanos não são “neutros” – e de serem inflexíveis, não moldáveis ao *sense-maker*, seus objetivos e necessidades, esses sistemas operam com base em previsões. Aquilo que aconteceu no passado é usado como material básico. Essas previsões negligenciam, por exemplo – porque algoritmos negligenciam – contingências corporais, história social e cultural, interações, possíveis mudanças e ajustes futuros, sendo cegos ao contexto, aos sentimentos e às emoções envolvidos e daí por diante. Não há, afinal, interação real acontecendo. Se *máquinas de fazer*

previsões não podem cooperar com seus usuários de forma equilibrada, e sustentar com eles uma relação, de fato transformando, crescendo e se modificando junto com eles; se bloqueiam as possibilidades de ajustes mútuos, elas evitam que emerja o novo, o surpreendente. As regras inseridas nos algoritmos tender a prevalecer. Pode-se dizer que a imprevisibilidade algorítmica tem ainda uma previsibilidade; ela acontece dentro de uma certa gama de regras e possibilidades.

Lidar com sistemas de ML, conseqüentemente, parece deixar pouco espaço para a surpresa e a flexibilidade que emergiriam das acomodações mútuas, trocas, conexões entre agentes cognitivos. Algo novo, nesse sentido, tende a surgir apenas como resultado dessas interações, isto é, como resultado de um genuíno contexto de *sense-making participativo*. Diante disso, se poderia questionar: menor possibilidade de surpresa leva a uma redução na *autonomia* ou simplesmente a uma redução na qualidade da experiência daquele cognoscente? De acordo com a lógica enativista, acredito que não seja possível separar esses dois aspectos; a dimensão do experimentar, do vivenciar parece ser uma camada significativa e intrínseca à *autonomia* do cognoscente¹⁵²; ligada aos conceitos enativistas de *emergência* e *experiência*. A qualidade da experiência está intimamente ligada também à afetividade, sendo as emoções uma dimensão essencial nos processos cognitivos humanos. Por que (e como) poderia um empobrecimento da qualidade da *experiência* vivida impactar a *autonomia* do cognoscente? Para entender plenamente esta parte do raciocínio, considero importante trazer de volta o conceito já mencionado de *adaptatividade* como desenvolvido pelo enativismo:

Uma entidade autopoietica pode ser capaz de lidar com perturbações sem a necessidade lógica de monitorar ativamente seu próprio estado e agir para melhorar as condições para a autopoiese continuada. Somente entidades autopoieticas adaptativas, que melhoram as condições para a continuidade da autopoiese, monitorando ativamente seu próprio estado, identificando pelo menos algumas tendências que as aproximam dos limites de viabilidade e neutralizando estas tendências, podem ser realmente "sense-makers" (Di Paolo, Rohde & De Jager, 2010, p. 50)¹⁵³.

¹⁵² A ideia da estética como uma dimensão incontornável de nossa atividade cognitiva é uma marca do trabalho do pragmático John Dewey – ver por exemplo Dewey (2008), e foi explorada por Mark Johnson (2017, 2018), bem conhecido por suas ideias a respeito da cognição corporificada. Johnson (2018, p. 2) nos define como "homo aestheticus – criaturas da carne, que vivem, pensam e agem em virtude das dimensões estéticas da experiência e da compreensão. Sobre essas conexões, ver também Leporace e Gondin, 2021.

¹⁵³ No original: *An autopoietic entity can be robust to perturbations without the logical necessity to actively monitor its own state and act to improve the conditions for continued autopoiesis. Only adaptive autopoietic entities that improve the conditions for continued autopoiesis, by actively*

Como Di Paolo e colegas dizem em Di Paolo, Cuffari e De Jaegher (2018, pp. 32-33): “[A]daptividade (Di Paolo, 2005) é o que permite aos corpos vivos distinguir uma situação como um risco ou uma oportunidade, para dizer a diferença entre bom ou melhor, mau ou pior”¹⁵⁴. Ao lidar com o mundo que o agente se põe a conhecer a partir de sua *experiência* – orgânica, sensório-motora e intersubjetiva – ele pode avaliar a viabilidade de suas ações e se deslocar para longe de situações que podem ameaçar sua *autonomia*, aproximando-se daquelas que podem garanti-la e potencializá-la.

[S]ense-making é a capacidade de um sistema autônomo de regular de forma adaptativa a sua operação e a sua relação com o meio, dependendo das consequências virtuais para a sua própria viabilidade como uma forma de vida. Ser um sense-maker implica em ter uma sintonia contínua (muitas vezes imperfeita e variável) com o mundo e uma prontidão para a ação. Através da combinação da auto-individuação material e precária e da regulação adaptativa das relações com o meio, a produção de sentido naturaliza o conceito de normas vitais e está no cerne de toda forma de ação, percepção, emoção e cognição, uma vez que em nenhuma instância destas está ausente a estrutura básica de preocupação e cuidado. Isso é o que distingue de maneira constitutiva a forma de vida mental de outros processos materiais e relacionais (Di Paolo, Cuffari & De Jaegher, 2018, p. 33)¹⁵⁵.

Quanto mais o *sense-maker* conhece o mundo, mais o experimenta em sua totalidade, estando imerso nas perturbações e desequilíbrios que lhe exigem adquirir habilidades diferentes para lidar com elas. Variados tipos de experiências parecem aumentar as possibilidades de desenvolver essas habilidades, já que recrutam uma diversidade de padrões sensório-motores, orgânicos e intersubjetivos, levando os *sense-makers* a tornarem-se *skillful copers*, como Dreyfus (2016) propõe. Parece então correto assumir que, por outro lado, reduzir as possibilidades

monitoring their own state, identifying at least some tendencies that bring them closer to the boundary of viability and counteracting these tendencies can be actual ‘sense-makers’ (Di Paolo, Rohde & De Jaegher, 2010, p. 50).

¹⁵⁴ No original: *[A]daptivity (Di Paolo, 2005) is what enables living bodies to distinguish a situation as a risk or an opportunity, to tell the difference between good or better, bad and worse. (Di Paolo, Cuffari e De Jaegher, 2018, pp. 32-33)*

¹⁵⁵ No original: *[S]ense-making is the capacity of an autonomous system to adaptively regulate its operation and its relation to the environment depending on the virtual consequences for its own viability as a form of life. Being a sense-maker implies an ongoing (often imperfect and variable) tuning to the world and a readiness for action. Through the combination of material and precarious self-individuation and adaptive regulation of the relations to the environment, sense-making naturalizes the concept of vital norms and lies at the core of every form of action, perception, emotion, and cognition, since in no instance of these is the basic structure of concern and caring ever absent. This is constitutively what distinguishes mental life from other material and relational processes (Di Paolo, Cuffari & De Jaegher, 2018, p. 33)*

de experimentar diferentes demandas pode resultar em uma redução das possibilidades de desenvolvimento dessas habilidades.

III. A falta de consciência acerca de riscos potenciais

Para além da necessidade de um *sense-maker* de resguardar a sua autonomia continuamente, agindo no mundo e aprimorando suas habilidades enquanto lida com o meio, o enativismo tem ainda algo importante a dizer sobre as normas que guiam a interação. Para se manter seguro e orientado a aquilo que melhor pode se adequar a ele, o cognoscente precisa estar ciente dessas normas, assim como dos riscos que pode enfrentar ao se acoplar com o mundo. O agente pode, ainda, colaborar para a revisão dessas normas. Mas como poderia o *sense-maker* ser capaz de decidir o que é melhor para sua vida e *autonomia* se lida com um sistema orientado por regras e premissas que desconhece e que não colaborou para definir? Ele seria capaz de fazer as melhores escolhas possíveis para a sua adaptatividade, nesse cenário?

Em primeiro lugar, dado que sistemas de machine learning são baseados em previsões, que são enraizadas no passado, eles tendem a limitar as possibilidades para o *sense-maker* – o que nos leva de volta à seção 4.2.2. Em segundo lugar, isso se agrava por conta do fato de que sistemas são geralmente orientados a alguns objetivos pré-definidos que, além de rígidos em comparação com a adaptatividade de um *sense-maker*, não são transparentes – o que nos leva de volta à seção 4.2.1.

A “normatividade” criada pela máquina é determinada pelos objetivos que se espera que ela atinja. Ela tende a repetir o passado, o que pode projetar possibilidades futuras reduzidas, assim criando um ambiente mais restrito ao cognoscente. Afinal, elementos que não fazem parte do passado capturado pelos algoritmos não poderiam fazer parte do futuro que eles podem prever. O problema é agravado porque quem usa o sistema nem sempre está ciente de que ele foi programado com certos objetivos e nem mesmo conhece esses objetivos. É frequentemente alardeado que sistemas de ML são capazes de adivinhar as necessidades dos usuários a partir de sua própria atividade, mas aqui exploro justamente a implausibilidade dessa espécie de mantra do marketing baseado na ciência de dados.

Mesmo que o cognoscente esteja ciente acerca de como esses sistemas funcionam, é geralmente difícil intuir quais objetivos são esses por trás das máquinas. Nesse sentido, esses sistemas são como caixas pretas, o que se agrava por não interagirem de fato conosco, não terem corpos, expressões ou emoções (nesse sentido, se parecem com a maneira como o cognitivismo concebe o outro numa interação; como uma caixa preta difícil de decifrar que recruta habilidades inferenciais para a resolução de problemas, como se fosse um quebra-cabeças complicado; sem aderência à troca fluida de experiência, afetividade e mútua comunicação/conexão que configura a conexão presente na intersubjetividade enativista).

Então, como essa situação (do agente que lida com um sistema de ML) poderia estar alinhada à preservação da autonomia tal como proposta no enativismo? Quando lidamos com outras pessoas, mesmo que seja uma situação desigual de força e de poder, podemos tentar restabelecer o equilíbrio aparentemente perdido, agindo para restaurar nossa autonomia por meio do diálogo, das trocas, da busca por uma reconexão. Podemos nos esforçar para manter ou para reassegurar o equilíbrio entre a nossa operação interna e o meio exterior, incluindo o outro; é possível manter o equilíbrio da interação enquanto mantemos nossa autonomia individual como cognoscentes. Mas, com um sistema algorítmico, não há troca real e esses recursos não funcionam como tentativa de preservação da autonomia do cognoscente.

A concepção enativista é orientada por uma ética fortemente calcada na relação com o outro e no cuidado que as trocas envolvidas exigem. Segundo Di Paolo, Cuffari e De Jaegher (2018, p. 310), essa ética, porém, não é um ingrediente adicionado posteriormente a corpos linguísticos já constituídos; ela é a própria “essência de cada ato linguístico”, seja orientado ao próprio ser, a outros corpos linguísticos ou ao mundo. A normatividade social e cultural também vai sendo tecida na medida em que corpos linguísticos se constituem. Pensar que essa seria uma camada adicionada posteriormente, somando-se a uma natureza original já constituída em separado, seria pensar nesses fatores a partir de uma perspectiva dualista, a qual o enativismo nega.

A ética, segundo o enativismo, é um conhecimento prático, um saber-fazer ético (*ethical know-how*, ver Di Paolo, Cuffari & De Jaegher, 2018, p. 313), que nos leva de volta ao conceito de *skillful coping* de Dreyfus, mas com uma forte

camada de intersubjetividade. Esse conhecimento ético-prático ou saber-fazer ético pode ser adquirido de uma maneira semelhante àquela como se aprende a andar de bicicleta, no sentido de que, ainda que se possa aprender teoricamente sobre as normas que regem uma comunidade por meio da compreensão de regras abstratas, uma aprendizagem mais profunda somente acontece por meio da prática. Também é na prática, nos atos, nas falas e nos gestos que a atitude ética (ou antiética) aparece: quando se declara amor por alguém, quando pessoas se unem em torno de uma causa comum, quando se demonstra solidariedade a alguém, quando um grupo é reconhecido ou rejeitado, quando se ensina algo a alguém, quando se cuida de alguém, e por aí vai (exemplos em Di Paolo, Cuffari & De Jaegher, 2018, p. 313).

É importante destacar que corpos linguísticos sempre podem questionar as normas que os têm regido, criticando as origens dessas normas, revendo os seus atos e entrando em diálogo sobre eles. Por conta disso é que destaco o aspecto da obscuridade das normas que regem os sistemas de machine learning e a complexidade que surge com relação a essas reflexões quando se trata de sistemas que não são *sense-makers*. Apenas corpos linguísticos, ou *sense-makers*, são capazes de discutir e rever as normas que compõem a esfera da ética que os rege. Quando essas normas não podem ser discutidas, elas passam a ser impositivas e aí, então, criam um dualismo com os seres cognoscentes que são regidos por elas – dualismo esse que, segundo lógica enativista, não deveria existir.

5.4

Estendendo a discussão: modelos matemáticos, fragmentos do mundo

Compreender por que sistemas algorítmicos podem ser tomados como caixas pretas fica mais intuitivo a partir da noção de modelo matemático, didaticamente explicada pela matemática Cathy O’Neil (2020).

Um modelo (...) nada mais é do que a representação abstrata de algum processo, seja um jogo de beisebol, a cadeia logística de uma petroleira, as ações de um governo estrangeiro, ou o público de um cinema. Esteja ele rodando dentro de um computador ou na nossa cabeça, o modelo pega o que sabemos e usa isso para prever respostas em situações variadas. (O’Neil, 2020, p. 30).

Como um exemplo simples, ela cita o modelo “informal” que usa todos os dias para cozinhar para a sua família.

As entradas ou inputs do meu modelo interno de culinária são as informações que tenho sobre minha família, os ingredientes que tenho às mãos e que posso conseguir, e minha própria energia, tempo e ambição. As saídas, ou outputs, são o que e como decido cozinhar. Avalio o sucesso de uma refeição pelo quão satisfeita minha família parece ao terminar, quanto comeram, e quão saudável a comida era. Ver se a refeição foi bem recebida e quanto foi aproveitada me permite atualizar o modelo para a próxima vez que cozinhar. As atualizações e ajustes fazem dele o que os estatísticos chamam de “modelo dinâmico. (O’Neil, 2020, p. 31).

O’Neil segue adiante e explica que, se quisesse formalizar o modelo, para que, por exemplo, pudesse facilmente explicá-lo para a mãe se tivesse que deixar os filhos com ela enquanto viajassem, poderia sistematizá-lo e torná-lo matemático. O ideal seria treinar esse modelo com o tempo, alimentando-o com dados diariamente sobre o que mudou e como foi a reação da família e daí por diante. Poderia incluir parâmetros, restrições, regras – por exemplo, fulano gosta de carne e fulano de pão ou macarrão. Ao fazer isso, ela teria transformado seu modelo interno informal em um modelo externo formal. “Ao criar meu modelo, estenderia meu poder e influência sobre o mundo. Estaria criando um eu automatizado que pode ser implementado por outros, mesmo que eu não esteja aqui”. (O’Neil, 2020, p. 33). Ela diz, ainda, que esse modelo, porém, sempre contaria com erros, “porque modelos são, por sua própria natureza, simplificações” (O’Neil, 2020, p. 33). Quanto a outras implicações acerca de um modelo, ela indica que:

- Um modelo sempre deixa alguma informação importante de fora (por exemplo, no modelo da comida, ela pode se esquecer de inserir a informação de que cenoura crua é mais apreciada do que cozida por sua família);

- Para criar um modelo, é preciso fazer escolhas sobre o que é importante para ser incluído, “simplificando o mundo numa versão de brinquedo que possa ser facilmente entendida, e a partir da qual possamos inferir fatos e ações importantes” (O’Neil, 2020, p. 33);

- Precisamos admitir que o modelo se presta a fazer um trabalho específico e poderá agir como “uma máquina ignorante com enormes pontos cegos” (O’Neil, 2020, p. 34); esses pontos cegos “refletem os julgamentos e prioridades de seus criadores” (O’Neil, 2020, p. 34);

- Os modelos têm a reputação de serem imparciais, mas “refletem objetivos e ideologias” (O’Neil, 2020, p. 35). Por exemplo, se ela tira os doces do seu modelo de alimentação para a família, com isso impõe essa restrição a ela. Essas escolhas

são resultado de valores e desejos pessoais. “Modelos são opiniões embutidas em matemática” (O’Neil, 2020, p. 35).

- Avaliar se um modelo funciona ou não é questão de opinião, já que a definição de sucesso de um modelo é um componente que faz parte do próprio modelo (e, portanto, também vem de quem o modela). “Em cada caso, devemos nos perguntar não somente quem desenhou o modelo, mas o que aquela pessoa ou empresa está tentando alcançar (...) Se meus filhos estivessem criando o modelo, sucesso seria sorvete todos os dias” (O’Neil, XX, p. 35);

- Um modelo fica obsoleto se não for atualizado constantemente;

- Modelos podem ser extremamente simples e funcionar – como no caso do alarme de incêndio que só precisa da variável fumaça para ser acionado, mas também podem exigir inúmeras variáveis.

Dois pontos são essenciais quanto à complexidade envolvida em modelos matemáticos, e ficam mais claros nos exemplos que O’Neil fornece. Um deles é o sobre a gravidade ou consequência dos chamados pontos cegos de determinados modelos. Isso significa que muitos atributos considerados essenciais podem ser deixados de fora quando não deveriam. Menos problemáticos seriam os pontos cegos no caso de modelos, por exemplo, de mapas. Prédios são ignorados quando pedimos ao Google Maps para nos mostrar o caminho, mas porque não são elementos relevantes a essa tarefa. Mas há casos em que pontos cegos se multiplicam e se complexificam, como no exemplo relacionado ao racismo:

Racismo, no nível individual, pode ser visto como um modelo de previsão zunindo em bilhões de mentes humanas ao redor do mundo. É construído a partir de dados incompletos, defeituosos ou generalizados. Quer venha de experiência própria ou de boatos, os dados indicam que certos tipos de pessoa se comportaram mal. Isso gera uma previsão binária de que todas as pessoas daquela raça irão se comportar da mesma forma. Desnecessário dizer, racistas não gastam muito tempo buscando dados confiáveis para treinar seus modelos deturpados (...) Ele é alimentado por coleta irregular de dados e correlações espúrias, reforçado por injustiças institucionais e contaminado por viés de confirmação. (O’Neil, 2020, p. 37)

Vejamos mais um exemplo em O’Neil (2020, p. 34), este relacionado ao campo da educação. Se uma escola decide avaliar professores com base nas notas de seus alunos, ignorando o quanto eles “envolvem e cativam, trabalham em tarefas específicas, lidam com o controle da sala ou os ajudam com problemas pessoais ou familiares” – do ponto de vista do gestor da instituição, o modelo seria capaz de fornecer “uma ferramenta efetiva de identificar centenas de supostos professores

insatisfatórios, mesmo sob o risco de alguns resultados falso-positivos” (O’Neil, 2020, p. 34). Por outro lado – ela coloca – seria uma simplificação que talvez sacrificasse a precisão, uma análise mais criteriosa, em troca de atender à eficiência.

Essa simplificação a que O’Neil se refere é justamente aquilo que pode ter um custo alto demais quando se trata de delegar tanto para os algoritmos quanto temos feito, enquanto sociedade automatizada. Simplificações de espaços urbanos na forma de mapas podem até facilitar as nossas vidas, pois selecionam o que importa para chegarmos de um ponto de origem a um destino sem elementos adicionais que possam nos distrair; por outro lado, os fatores que são importantes quando se trata de avaliar professores ou alunos não são tão simples assim de eleger. Dependem do “modelo” que a escola adota, daquilo que valoriza e acha relevante na maneira de o professor atuar, da forma como avalia o desempenho dos alunos, dos interesses de várias partes a que pretende atender. Depois que esse modelo é desenhado, se ele é transposto para um sistema de machine learning, passa a ser alimentado por dados que são gerados justamente para alimentá-lo. Um dos maiores problemas é que os critérios selecionados para compor um modelo não são transparentes. As pessoas cujos dados são considerados nem mesmo sabem que eles estão servindo a determinado fim, ou como estão sendo combinados e usados. Com esse círculo estabelecido, os pontos cegos se fortalecem ainda mais, pois quaisquer casos novos que surjam não entram nesse modelo, e ele pode fortalecer cada vez mais os vieses. Alguns casos serão mencionados no capítulo a seguir. De todo modo, cabe salientar que neste trabalho o esforço é pela compreensão dos pressupostos, por isso o foco em casos específicos é reduzido.

5.5

Conclusão do capítulo

A partir do esforço empreendido aqui, e aplicando conceitos do enativismo a machine learning, foi possível concluir que sistemas algorítmicos não são *sense-makers*. Se fossem, deveriam ter uma permeabilidade ao ambiente externo de um modo que pudessem estar continuamente participando de uma interação corporificada, intersubjetiva, flexível e moldável pelos dois agentes envolvidos, de uma maneira tal que ambos estivessem cientes das normas vigentes e assim pudessem fazer movimentos para manter sua autonomia no melhor nível possível.

O sistema composto pelos dois agentes teria uma autonomia própria, mantida pela autonomia dos outros dois individualmente, e manteria as suas trocas com o meio externo também com vistas a orientar-se para se manter autônomo. Por outro lado, isso não acontece justamente porque sistemas artificiais não têm os requisitos para entrar nessa interação em que se observa *sense-making participativo*. As definições de *sense-making participativo* não se aplicam quando se lida com sistemas artificiais algorítmicos. Os sistemas inteligentes não são inteligentes quando se trata de emoções e afetividade. Reciprocidade seria um fator fundamental em qualquer relação, em qualquer interação, e não se pode esperar isso ao lidar com um sistema algorítmico que segue determinado modelo.

Como demonstrado, seres cognoscentes, quando em equilíbrio, buscam não apenas um acoplamento fluido e contínuo com o ambiente para a manutenção de sua autonomia, mas *o melhor acoplamento possível*. Nessa “dança” com o mundo, aprendem sobre as normas que guiam o universo em que eles estão inseridos, de modo que possam agir de acordo com elas, contribuindo para modificá-las em alguma medida. Por outro lado, as normas subjacentes a sistemas de ML, ou a vários desses sistemas conectados entre si, é fixa e estabelecida externamente, o que a torna, em geral, desconhecida dos cognoscentes (que utilizam esses sistemas). Isso pode ameaçar as suas possibilidades de ação, uma vez que o engajamento ao qual ele está habituado – e que é regulado por ações corporais mútuas e intersubjetivas – permite que as experiências se destaquem pelo aspecto da surpresa, da imprevisibilidade e da constante adaptatividade oriunda do fato de estarmos sempre lidando com o ambiente e com as outras pessoas.

A análise elaborada neste capítulo se apoia em dois pilares: o da investigação do sentido da aprendizagem de máquina, procurando demonstrar razões por que ela não envolve experiência/percepção (*sense-making*), e o da investigação da interação que se faz possível entre seres humanos e tais sistemas, procurando demonstrar as razões pelas quais ela não poderia ser considerada interação de fato, ou *sense-making participativo*. Este aspecto da interação – e sua caracterização como social ou não – tem relevância não apenas porque esta tese deseja contribuir para a análise dos pressupostos subjacentes aos sistemas de machine learning, colaborando para que se entenda como poderiam ser usados na educação, mas porque o aspecto da participação coletiva na percepção do mundo, fazendo emergir sentido e significado, é um ponto crucial ao enativismo. A ética enativista é

construída justamente a partir do outro. A intersubjetividade, porém, não é apenas um ingrediente relacional quando se trata da mente humana, mas um aspecto constitutivo da cognição – de acordo com o olhar enativista.

No próximo capítulo, a ideia é explicitar os pressupostos para um conceito de aprendizagem que emerge com o enativismo e explorar indicações de caminhos para que se possa encontrar meios de inserir machine learning na aprendizagem mantendo a educação em linha com os pressupostos enativistas. Dois elementos em particular são explorados: o da corporificação e o da intersubjetividade. Para além da aprendizagem de máquina, seres humanos aprendem em cooperação uns com os outros e constroem sentido coletivamente. As tecnologias, digitais ou não, encontram-se nesse contexto. Vou procurar apresentar como investigações e iniciativas que procuram equalizar essas demandas podem complementar uma análise crítica de machine learning para fins de uma aprendizagem enativista.

In his famous novel 1984, George Orwell got one thing wrong. Big Brother is not watching you, he's watching us. (Eubanks, 2018)

No capítulo anterior, procurei aplicar o conceito de autonomia tal como elaborado por enativistas para compreender a relação humana com sistemas algorítmicos de machine learning. Quando o que se pretende é incluí-los na aprendizagem, essa análise conclui que há riscos – subjacentes à maneira como esses sistemas funcionam – à autonomia do cognoscente, no sentido de que esses sistemas não são autônomos, e com isso não promovem junto ao cognoscente uma relação de produção compartilhada de sentido como ocorre num contexto de *participatory sense-making*. Isso não significa, porém, que tais sistemas não possam participar de processos de aprendizagem que se pretendam enativistas. Se, por um lado, não é difícil dizer que sistemas artificiais estão longe de “agir” como humanos, por outro é necessário que se analise isso de modo mais profundo para descobrir as razões pelas quais eles são diferentes de nós e (n)o que isso pode impactar. As potencialidades do nosso encontro com tais tecnologias podem estar, justamente, nessas fissuras.

Isso fica bastante claro quando entendemos que esses sistemas não são agentes cognoscentes, pelas razões advindas do enativismo, explicitadas. O que essa análise mostra, então, revela limites que devem ser levados em consideração, no intuito de não confundir o que se pode delegar a esses sistemas, por exemplo. Revela, ainda, alguns pressupostos que ajudam a compreender como o enativismo perspectiva a aprendizagem. Há algo de frutífero e inesperado que é típico da intersubjetividade, já que agentes cognoscentes que agem juntos podem coparticipar numa produção de sentido e fazer emergir algo novo, que não existiria sem essa troca. Pode-se, então, concluir que um caminho coerente com uma aprendizagem enativista promove a um lugar de destaque as trocas intersubjetivas, que por sua vez têm como premissa a experiência enraizada no corpo.

Ainda não se sabe exatamente como sistemas de machine learning e sistemas algorítmicos em geral podem colaborar com a aprendizagem humana. Suas potencialidades e limites vêm sendo observados na medida em que esses sistemas vêm sendo utilizados. Mas, ao conduzir o olhar para as premissas subjacentes à

aprendizagem em machine learning, contrastando-as com a concepção de aprendizagem que pode emergir do enativismo, este trabalho de pesquisa tem a expectativa de contribuir com elementos que ajudem a orientar o debate acerca de como essas tecnologias podem estar presentes nos processos de aprendizagem humanos. Para tanto, no capítulo anterior, assinalei um mecanismo presente na lógica algorítmica que engendra riscos à autonomia do agente cognoscente, sob a perspectiva do enativismo. Pode ser pertinente, agora, perguntar: seria possível contornar esses riscos de alguma maneira, em vez de aceitá-los, simplesmente? Isto é, seria possível evitar que a imersão em um contexto permeado por algoritmos traga ameaças à autonomia do cognoscente? Possivelmente não há uma resposta certa para isso, muito menos uma resposta única. Mas acredito que o enativismo, ao mesmo tempo em que nos ajuda a enxergar riscos, pode também indicar caminhos para que determinados elementos sejam valorizados na aprendizagem, sendo essa valorização, em si, um caminho para a reflexão e busca de soluções.

Esses caminhos passam pelo esforço de privilegiar o corpo e o movimento na aprendizagem. Se há cada vez menos dúvida quanto à importância de se levar esses aspectos em consideração, a necessidade cresce ainda mais quando se habita um contexto em que calcular e quantificar estão na ordem do dia. Esse cenário promove uma redução da subjetividade a elementos em tese mensuráveis, precisos, mas que na realidade deixam escapar tantos aspectos essenciais de um indivíduo. Trata-se de uma rota que se mostra exatamente oposta à do discurso da personalização, da individualização, podendo representar uma tendência à massificação em detrimento de uma atenção a características particulares. Evidenciar o lugar do corpo ou, mais precisamente, dos corpos é assumir a premente necessidade de irmos além da economia de dados, ou de mudarmos os pressupostos dessa economia, buscando nossas raízes, buscando aquilo que nos torna exatamente quem somos e quem ainda podemos ser – aspecto essencial para sairmos dessa massificação e de uma circularidade que ameaça nos reduzir. O nosso porvir corporificado, o nosso *vir-a-ser* é, em si, uma resistência à *datafication*, à automatização cega de processos. Como inspiração, vale olhar para alguns trabalhos empíricos que exploram a presença ativa do corpo na aprendizagem. Trago alguns deles neste capítulo.

Conforme mencionado, além do corpo e suas subjetividades, a intersubjetividade é outro elemento que é caro ao enativismo – e que pode indicar caminhos para a manutenção da autonomia frente ao uso de sistemas algorítmicos.

Como argumentado no capítulo anterior, é justamente a troca intersubjetiva que falta quando lidamos com sistemas artificiais; a potência das trocas humanas fica ainda mais visível quando notamos os momentos em que elas faltam. Pode-se pensar que, se a mente humana somente existe a partir da intersubjetividade – a partir de uma perspectiva enativista – a aprendizagem é também produto dessas trocas. Há vários desdobramentos para isso. Um deles é que a educação não pode negligenciar o papel que as trocas entre indivíduos têm para a aprendizagem de qualquer tipo e em qualquer espaço, e deve considerá-las centrais, fundamentais. Temos acesso a tecnologias, algorítmicas ou não, não reduz a necessidade de trocas intersubjetivas; pelo contrário, apenas aumenta a sua importância. Até para afinarmos as nossas percepções quanto a aquilo que esses sistemas têm para nos oferecer, precisamos analisar profundamente esses efeitos. As análises e os seus resultados são um trabalho conjunto. Decorre disso que somente emergirão soluções se nós tivermos um empenho coletivo: este é um dos pontos, não o único, relativos à intersubjetividade nesse contexto.

Antes ainda de prosseguir, cabe ressaltar que aquilo que não pode ser capturado pelos sistemas artificiais na forma de dados, estatística, previsões, modelos matemáticos e algoritmos pode ser justamente o que há de mais desafiador quando se trata de aprendizagem. Isso é algo que se torna particularmente relevante se houver uma real preocupação em rever as premissas subjacentes à aprendizagem que ainda persistem. Pode-se dizer que os pontos cegos desses sistemas escondem aspectos que não deveriam ser negligenciados. De todo modo, buscar alternativas, olhar nas entrelinhas, gerar caminhos novos é algo típico dos *linguistic bodies* que somos. Somos autopoieticos e somos capazes de nos adaptar. Temos a nossa identidade, mas também a nossa flexibilidade. Temos bem demarcadas as fronteiras da nossa individuação, mas somos permeáveis a transformações, a um contínuo *vir-a-ser* ou *tornar-se*. Na medida em que buscamos a melhor situação possível para a nossa autonomia, individualmente como *sense-makers* e enquanto sociedade, como corpos linguísticos em *participatory sense-making*, aumentamos as possibilidades, podemos reduzir alguns riscos e modular a relação que temos com o ambiente em torno de nós, para que ela fique mais equilibrada. Este é o tema deste capítulo final. A pergunta que o impulsiona é: quais elementos advindos do enativismo podem ser empregados para enfrentarmos as ameaças à autonomia do cognoscente que podem aparecer em meio à algoritmosfera?

Para iniciar o capítulo, apresento uma síntese dos pressupostos para uma concepção de aprendizagem enativista. Apesar de o enativismo não fazer uma proposta automaticamente aplicável à educação, tampouco propor uma teoria da aprendizagem, ou justamente *porque* não faz isso, o esforço que aqui empreendo é justamente nesta direção: o de contribuir para ressaltar os elementos que podem ajudar nessa transposição. O enativismo é capaz de inspirar maneiras de ensinar que consideram a versão mais forte da participação do corpo, do ambiente, da afetividade e da intersubjetividade na aprendizagem. Para ilustrar, busquei alguns exemplos, que apresento no decorrer do capítulo. Depois, parto para as outras etapas da argumentação, abordando como a intersubjetividade representa tanto um marcador das diferenças entre a interação humano-humano e a “interação” humano-máquina (que deve aqui ser considerada entre aspas) como um fio condutor para que, juntos, possamos pensar criticamente sobre a nossa relação com esses sistemas. É olhando para esse elemento – constitutivo da mente humana, existindo um “nós” antes mesmo de um “eu” quando se trata de um agente cognitivo (Bannell, Leporace & Santos, 2021) – que se pode analisar o papel de sistemas de machine learning na educação.

6.1

Síntese de uma concepção enativista para a aprendizagem

A partir da proposta do enativismo para a cognição, a seguir apresento alguns dos principais pressupostos que considero estarem subjacentes a uma perspectiva da aprendizagem que emerge com o enativismo. A maioria desses pontos já apareceu no decorrer da tese. A ideia, agora, é sintetizá-los de maneira a esquematizar tais ideias e contribuir para encaminhar futuros estudos que possam desenvolver uma teoria de aprendizagem fundamentada no enativismo.

1. A aprendizagem é corporificada. Não é um processo computacional e representacional que se limita ao cérebro, mas demanda o corpo todo, em suas dimensões orgânica, sensório-motora e intersubjetiva, em acoplamento com o ambiente;

2. *A aprendizagem está diretamente ligada à afetividade. Aprender é afetar-se pelo mundo e afetar o mundo e o outro;*

3. *A aprendizagem acontece na precariedade. O ambiente com o qual interagimos é instável e oferece, justamente por isso, a oportunidade de nos mantermos em ajustes constantes que nos mantêm em aprendizagem constante;*

4. *A aprendizagem envolve processos de aquisição e aprimoramento de habilidades de vários tipos, que acontecem simultaneamente, assim como percepção e ação acontecem simultaneamente;*

5. *Nem todo processo de aprendizagem tem resultados programáveis ou metas que possam ser estabelecidas previamente;*

6. *A aprendizagem é sempre contextual. Como seres situados, sense-makers, aprendemos orientados ao mundo e a partir da nossa experiência de mundo;*

7. *A aprendizagem é um movimento contínuo. Estamos continuamente aprendendo, mesmo quando adultos, até o fim da vida;*

8. *É possível perceber avanços na aprendizagem, mas não é possível medir da mesma maneira todas as aprendizagens ou tipos de aprendizagem pelos quais o ser humano passa continuamente no decorrer da sua vida.*

9. *A aprendizagem é uma experiência estética em si. Estética não é uma categoria ou um tipo de experiência¹⁵⁶, mas uma dimensão da aprendizagem, portanto dela indissociável e que sempre está presente quando aprendemos;*

10. ***Aprendizagem é intersubjetividade.*** *A aprendizagem envolve produção de sentido não apenas no nível individual, mas no nível coletivo,*

¹⁵⁶ Ideias trabalhadas em Leporace e Gondim, 2021.

concomitantemente. Aprendemos com outros, e essas trocas não cessam, já que a aprendizagem é contínua e entrelaçada com o ambiente.

Tendo em consideração então que esses elementos são subjacentes a toda educação que se pretende enativista, sigo adiante a partir de dois elementos essenciais, a corporificação e a intersubjetividade, explorando demais pressupostos a partir desses dois, que considero centrais.

6.2

Da centralidade do corpo na aprendizagem

Por mais que a educação venha valorizando o corpo, considerando a sua importância para os processos de ensino e de aprendizagem, há ainda muita hesitação quanto à maneira como isso pode ser posto em prática no cotidiano. Na educação formal, é comum se considerar que o corpo deve participar das atividades com crianças pequenas, mas tende a ser minimizada a lembrança de que nunca paramos de aprender e sempre aprendemos com todo o corpo, mesmo adultos. Os pressupostos enativistas não se restringem a uma perspectiva infantil, apesar de serem úteis também para pensar a criança, claro. Não à toa, Mark Johnson (2007) – um dos que argumentam pela cognição corporificada – refere-se a adultos como “*big babies*”, ou bebês grandes/crescidos, no sentido de que podemos estender a necessidade de uma imersão corpórea no mundo para seguirmos aprendendo enquanto adultos.

Quando há tecnologias avançadas sendo aplicadas para fins educacionais, diversos outros aspectos importantes acabam deixados de lado – por exemplo, pode-se valorizar mais a usabilidade de um recurso do que os resultados de aprendizagem obtidos com ele, como detectaram Radianti et al. (2020, apud Johnson-Glenberg, 2022). Por outro lado, pesquisas focadas nos possíveis benefícios da inclusão do corpo na aprendizagem revelam que considerar os gestos e a corporificação de metáforas, por exemplo, pode trazer benefícios para aprendizes de várias idades e em diversos tipos de situações de aprendizagem, sejam formais ou informais.

Uma hipótese é que aprendizes engajados em altos níveis de corporificação irão aprender conteúdo mais rapidamente e de uma maneira mais profunda, porque ativar códigos sensorio-motores fortalece os caminhos da memória. Além disso, gestos

podem atenuar a carga cognitiva (Goldin-Meadow, 2011, 2014, apud Johnson-Glenberg, 2022)

Johnson (2021), que tem um trabalho bastante aprofundado sobre a relação entre os movimentos corporais e as metáforas, explica que “projeções a partir de partes do corpo (*body-part projections*) são significativas porque colocam em ação (*enact*) aspectos de nossas maneiras fundamentais de nos relacionar com nosso ambiente e de agir nele” (Johnson, 2021, p. 110). Junto a Lakoff (1987) ele desenvolveu um conceito chamado de “esquema-imagem”,

A ideia básica era a de que, dada a natureza de nossos corpos (como percebemos, como nos movemos, o que valorizamos) e as dimensões gerais de nossos meios (estruturas estáveis no ambiente), nós vamos experimentar padrões regulares recorrentes (como em cima-embaixo, direita-esquerda, em frente-atrás, contenção, iteração, equilíbrio, perda de equilíbrio, origem-caminho-meta, movimento forçado, locomoção, centro-periferia, reto, curvo etc.) que nos permitem (*afford*) interações significativas, tanto físicas como sociais, com nossos meios. (Johnson, 2021, p. 110).

O autor explica que, à medida que experimentamos sensações a partir da maneira como lidamos com objetos, e utilizando para isso nossos corpos, nos tornamos capazes de estender essas percepções a domínios mais abstratos – ou seja, *metafóricos*:

O fato de que humanos existem e operam dentro do campo gravitacional da Terra gera experiências repetidas de relações para cima/para baixo (*up/down*; isto é, *verticais*). Nós compreendemos objetos como subindo (*rising up*) e caindo (*falling down*), como estando na posição vertical (*upright*) ou horizontal (*lying down*), assim como em cima de ou embaixo, em relação à nossa orientação corporal e às nossas experiências. O fato de que nós, rotineira e crucialmente, experienciamos o equilíbrio ou a falta de equilíbrio dá origem a um esquema de equilíbrio que se aplica literalmente a objetos físicos que se equilibram e, metaforicamente, aos nossos estados mentais internos, a equações matemáticas, e a noções de justiça e equidade política (Johnson, 1987). Nossas várias experiências diárias com recipientes e espaços contidos geram um esquema constante de recipiente em nosso sistema de percepção (Lakoff, 1987). Nossos vários encontros diários com objetos que se movem e com os movimentos de nossos próprios corpos dão origem a um esquema de locomoção (Dodge & Lakoff, 2005). (Johnson, 2021, p. 110).

Outro aspecto importante dessas percepções e interações com os objetos do mundo é que elas são multimodais; isso se alinha com a ideia enativista de que a percepção do mundo se dá em diversos níveis ao mesmo tempo – são níveis de *sense-making*, com a aplicação de diversos recursos corporais diferentes. É interessante observar um exemplo oferecido por Johnson (2021), abaixo, para comparar a maneira como um ser humano pode absorver um conceito, como o de cadeira, em comparação com um sistema de machine learning:

Por exemplo, entender um conceito como cadeira envolve uma simulação sensorial, motora e afetiva de possíveis experiências com cadeiras de todos os tipos. Tais simulações irão envolver múltiplas modalidades (como visão, toque, audição e propriocepção), na medida em que nossas interações com cadeiras são multimodais. Vemos cadeiras sob vários pontos de vista quando andamos em volta delas, sabemos como é sentar em vários tipos de cadeiras feitas de diferentes materiais e tocá-las, e sabemos quais os tipos de programas motores exigidos para sentar e levantar de cadeiras. Saber o significado de cadeira, entender o que uma cadeira é em certo contexto, é simular experiências com cadeiras usando todas as modalidades sensoriais, motoras e afetivas disponíveis para nós. (Johnson, 2021 p. 110)

Podemos, enquanto seres cognoscentes corporificados, *sentir* o que é uma cadeira, perceber e compreender com a nossa atividade corporal o que é uma cadeira; nos utilizaremos dessa experiência na próxima vez em que tivermos que “decidir” se algo que vemos ou sentimos é de fato uma cadeira em que podemos nos sentar ou não. Enquanto isso, um sistema de aprendizagem de máquina precisaria ser “apresentado” a uma enorme variedade de imagens desse objeto para que “aprendesse” o que é uma cadeira. Mesmo assim, poderia falhar diante de um exemplo mais ousado e que saísse do esperado, já que cadeiras não são sempre iguais e pode haver um designer a projetar uma que fuja completamente ao padrão mais conhecido. Claro, em todos os casos, para que se diga que uma cadeira é uma cadeira, algo na organização dela precisaria ser preservado; caso contrário, a essência do objeto se perderia e ele não seria mais uma cadeira (essa é uma ideia presente no enativismo, encapsulada no conceito de clausura operacional; isto é, a organização de um sistema precisa ser mantida para que ele se mantenha individuado e mantenha a sua identidade; no caso de sistemas vivos, a organização é autopoietica).

Exercitando mais um pouco e saindo desse exemplo da cadeira para estendê-lo a uma outra situação, imaginemos que um sistema artificial baseado em redes neurais seja treinado para classificar imagens do clima de regiões do Brasil por meio de imagens cadastradas. Após ter acesso a grandes quantidades de imagens, o sistema “aprende” a classificá-las e se torna um grande “conhecedor” do clima brasileiro. Esse é um dos casos em que, ironicamente, uma máquina se mostraria uma conhecedora muito mais potente do que nós. Afinal, um sistema de machine learning é capaz de “conhecer” todos os sabores na teoria. Porém, vale lembrar, não conhece nenhum de fato. Também pode-se ter uma rede neural que “conheça” todos os climas ou a vegetação do mundo inteiro, sem nunca ter visto e nem sentido nenhum deles. Se fôssemos ensinar o clima de uma determinada região do mapa

geográfico, poderíamos, de modo semelhante, mostrar imagens às crianças de uma classe, esperando que elas compreendessem e reconhecessem características do clima tropical, árido, semiárido, o que é a neve ou uma neblina etc. Mas, se essas crianças nunca estiveram em uma região árida ou semiárida, por exemplo, elas somente poderão ter a compreensão teórica do que aquilo pode significar. Assim como, se nunca tocaram num floco de neve, não têm a experiência de sentir o gelado da neve e sua textura, sua cor ou densidade. Se jamais estiveram num local em que o céu estivesse cheio de neblina, não sabem como é perder a noção do horizonte na paisagem diante de si. Claro, a percepção pode ser intensificada por meio de vídeos ou até de uma “visita” virtual por meio do *Google Street View*, mas esses recursos não podem substituir as sensações proporcionadas pela experimentação que envolve o corpo no mundo, o corpo em contato com o mundo, sentindo um clima seco ou úmido, frio ou quente, enfim. Por outro lado, se tivessem a vivência física de determinado clima, os alunos não precisariam de milhares de imagens para aprender sobre aquele clima. A experiência do sentir, do vivenciar, presenciar, seria suficiente. Sob a perspectiva enativista, aprender de fato é algo que recruta essa vivência, até mesmo para que se tenha noção daquilo que se conhece de verdade e aquilo que ainda está para ser conhecido. Somos capazes de diferenciar o que conhecemos de fato daquilo que apenas conhecemos em teoria; por exemplo, um sabor que nunca provamos, que sabemos que existe, mas que para nós é considerado desconhecido porque ainda não experimentamos; um lugar ou uma pessoa que apenas conhecemos por fotos.

Isso pode ir muito mais longe: a experiência não é a marca apenas do instantâneo, do sabor sentido, mas daquilo que se vive e assim modifica o indivíduo nos níveis orgânico, sensório-motor e intersubjetivo (mas não modifica um sistema artificial, que, justamente, não tem uma experiência corporal nesses níveis). Poderíamos conceber que, por exemplo, uma determinada pessoa que nasceu numa cidade do interior com cinquenta ou cem mil habitantes e nunca morou numa cidade movimentada como o Rio de Janeiro ou Nova York, nem mesmo visitou uma cidade assim, teria dificuldades em realmente compreender o que esse adjetivo “movimentada” significa; mesmo que visitasse uma cidade assim, morar seria diferente e modificaria as percepções dessa pessoa de uma forma diferente. Mudam as concepções de distância – o que seria “perto” ou “longe; mudam as noções de tempo – “rápido” ou “demorado”; mudam as percepções do que seria um local cheio

de gente ou vazio, as sensações de segurança ou perigo, e daí por diante. A compreensão desses aspectos se daria pela experimentação diária de caminhar pela cidade ou utilizar um transporte para se deslocar, trabalhar ou estudar na cidade, se relacionar com pessoas, resolver problemas, desenvolver atividades de lazer, enfim, todos os aspectos de uma vida cotidiana.

Retornando aos exemplos do cotidiano escolar, um recurso de aprendizagem simples e que vem mostrando resultados interessantes em pesquisas é o da utilização do próprio corpo dos alunos como meio – ainda que não saiam da sala de aula ou de um espaço de aprendizagem restrito. Segundo Boaler (2022), estudos em neurociência e em comportamento revelam conexões entre avanços na percepção dos dedos (*finger perception*) e aprendizagem de matemática. A percepção dos dedos é a capacidade de diferenciar os dedos, na qual, segundo neurocientistas, escolas deveriam investir mais (Boaler, 2022, p. 122-123).

Os pesquisadores descobriram que quando crianças de seis anos de idade melhoraram a qualidade da percepção dos dedos, melhoraram no conhecimento de aritmética, particularmente a subitização (a capacidade de estimar uma quantidade em um conjunto sem contar cada item), a contagem e a ordenação numérica (Boaler, 2022, p. 122)¹⁵⁷

Interessante também é um estudo de pesquisadores de Stanford que constatou que estudantes do primeiro ano que utilizaram um dispositivo robótico tiveram melhor resultado do que aqueles que tiveram acesso apenas a um computador para resolver os mesmos tipos de problemas. O dispositivo robótico ajudava a desenvolver a percepção dos dedos, já que os estudantes usavam os dedos para escolher respostas para problemas matemáticos e recebiam uma resposta tácita diretamente neles. (Martinez et al., no prelo, apud Boaler, 2022). Outros estudos mostram a importância do posicionamento do corpo na aprendizagem de conceitos. Para aprender sobre ângulos, por exemplo, alunos que posicionaram seus corpos de uma determinada maneira demonstraram obter um resultado melhor do que outros que foram apresentados aos conceitos sobre ângulos por meio de uma instrução tradicional, segundo Gallagher e Lindgren (2021). Ganhos de aprendizagem foram percebidos ainda em estudantes que “encenaram trajetórias celestes com seus

¹⁵⁷ No original: *Researchers found that when six-year-old children improved the quality of their finger perception, they improved in arithmetic knowledge, particularly subitizing (the ability to recognize a number in a set without counting), counting and number ordering. (Boaler, 2022, p. 122)*

corpos ou objetos representando estrelas e planetas” (Gallagher & Lindgren, 2021, p. 154).

Gallagher e Lindgren (2021) indicam, ainda, um exemplo em que se pode observar uma situação de aprendizagem envolvendo metáforas enativas e *sense-making participativo*. Trata-se da brincadeira de faz-de-conta, em que, no caso observado, uma criança pega uma banana e a aproxima da orelha como se fosse um telefone, após observar sua cuidadora fazer o mesmo. A leitura enativista dessa situação nos indica que se trata do exercício de uma habilidade motora básica “motivada por um processo intersubjetivo de imitação” (Gallagher & Lindgren, 2021, p. 151).

Isto é parte de uma interação corporificada que o bebê capta imediatamente. O bebê vê a banana como um telefone, não por meio de um conjunto de representações mentais, mas nas ações da cuidadora, ao captá-las e ao imitá-las – um exemplo claro de *sense-making participativo*. (Gallagher & Lindgren, 2021, p. 151)

Os autores também destacam as diferenças que se poderia observar caso se adotasse uma perspectiva cognitivista para a brincadeira em questão.

Na perspectiva excessivamente cognitivista, a transformação imaginativa é simplesmente “substituir um conteúdo de pensamento por outro”, porém, “acessando e controlando inputs (crenças e desejos) para as ações de projeção imaginativa que sustentam a simulação” (Currie & Ravenscroft, 2002: 140). Essa interpretação baseia-se fortemente em estados semelhantes a crenças e em processos de pensamento a sustentar tais habilidades. Currie, então, intelectualiza demais o processo. “Na simulação, se age a partir de uma suposição, por exemplo, de que a caixa em que estou sentando é um carro; na suposição, pode-se considerar uma ideia, extrair consequências a partir dela, considerar evidências para isso, e compará-la com outras ideias” (Currie, 2003: 233). Não é claro, no entanto, que a criança esteja fazendo nada disso enquanto ativamente segura a banana como um fone ou dirige a caixa como se fosse um carro. (Gallagher & Lindgren, 2021, p. 151).

Ainda neste tópico, vale observar o trabalho que esses mesmos autores fizeram para investigar empiricamente os impactos de recursos tecnológicos digitais na aprendizagem envolvendo metáforas enativas. O projeto MEteor é descrito a seguir:

Crianças interagem com o MEteor usando seus corpos para lançar um asteroide a uma certa velocidade e então prever para onde ele irá se mover com base na presença de planetas com forças associadas. Isso envolve uma metáfora enativa em que a criança se identifica com o asteroide – “Eu sou o asteroide” – e então põe em ação o comportamento do asteroide. Ao usarem o MEteor, as crianças são guiadas por uma série de níveis em que encontram uma progressão de ideias críticas e princípios da física sobre como objetos se movem no espaço, como, por exemplo, a aceleração gravitacional e as Leis de Kepler para o movimento planetário. O feedback acerca

da precisão das previsões das crianças é entregue dentro do quadro da metáfora, em simulações em tempo real que solicitam que a criança incline sua trajetória de maneira mais brusca em torno de um planeta para, de forma precisa, retratar os efeitos das forças gravitacionais. As crianças são capazes de construir seu entendimento em torno dos movimentos de seus próprios corpos, com suportes representacionais (gráficos e outras visualizações) construídos no ambiente de uma maneira que dá apoio à construção desse novo conhecimento. (Gallagher & Lindgren, 2021, p. 156)

Duas versões de interação foram utilizadas na experiência, que apareceu primeiro em Lindgren e Moshell (2011) e Lindgren et al. (2015). Uma foi considerada a versão enativa fraca e a outra, a versão forte. Na fraca, os participantes usaram uma versão para *desktop* do programa MEteor, controlando o asteroide pelos movimentos do mouse. Já na versão forte da experiência, os participantes usaram o corpo todo para controlar o asteroide, imersos em um espaço interativo do tamanho de uma sala, feito para que eles se sentissem “dentro” do sistema sobre o qual estavam aprendendo (Gallagher & Lindgren, 2021, p. 157). Os resultados são descritos a seguir:

Lindgren e Moshell (2011) descobriram que os participantes (62 estudantes do ensino fundamental II divididos pelas duas condições) do grupo enativo forte construíram “diagramas de aprendizagem” depois de usar a simulação, que incluíam representações mais dinâmicas (por exemplo, setas); eles se mostraram menos dependentes das características da superfície da simulação (por exemplo, estrelas no cenário de fundo e a textura dos planetas) em comparação com os participantes da condição enativa fraca. As descobertas sugerem que participar de metáforas enativas leva a uma compreensão mais abrangente e flexível do domínio. Em Lindgren et al. (2015), dos 113 participantes do ensino fundamental II, 58 usaram a versão enativa forte da simulação, e 55 usaram a versão fraca. Resultados mostraram que estudantes na condição forte pontuaram mais alto nas avaliações padronizadas de conhecimentos em física. Essas avaliações incluíam não somente questões sobre como objetos se movem no espaço, mas também questões que endereçaram o entendimento da física de maneira mais ampla (por exemplo, os efeitos da gravidade em objetos lançados no ar da superfície da Terra). As pontuações mais altas nessas questões mais gerais sugerem que metáforas enativas podem ter impactos profundos e transversais no raciocínio dos estudantes, que se transferem para além do contexto específico da simulação interativa. (Gallagher & Lindgren, 2021, p. 157).

Além desses resultados, os autores (2021, p. 158) observaram nos alunos efeitos sobre a disposição para aprender. Alguns reportaram concordar mais com frases como “Eu gosto de falar com outras pessoas sobre ciência”, descoberta que “sugere que metáforas enativas habilitadas por tecnologias podem dar aos estudantes mais sensações de controle e confiança em seu próprio aprendizado”. Essa constatação indica que houve uma transformação nesses alunos após a

vivência de tais processos de aprendizagem, não apenas uma absorção conteúdo pretendido. Mais adiante, voltarei a falar de aprendizagem e transformação.

6.2.1

Corpo, transformação e percepção da aprendizagem

Todos esses exemplos de como o corpo pode estar mais presente na educação, inclusive quando se trata de trazer tecnologias para a aprendizagem, contrastam com a rigidez algorítmica dos modelos matemáticos e do *big data*. O que os dados significam? O que há para além dos dados? Essas são perguntas que precisam estar presentes quando se trata de inserir machine learning no contexto da aprendizagem humana. Fazer perguntas sobre dados, sobre o que podem significar e o que possivelmente *não* englobam, é essencial. Especialistas no tema, como Coeckelberg (2020, por exemplo), O’Neil (2016, por exemplo) e Véliz (2020, por exemplo), concordam com a ideia de que o fator humano é de extrema importância nos processos de tomada de decisão envolvendo algoritmos. Os impactos podem ser ainda maiores quando se trata de educação, já que os efeitos são sentidos no desenvolvimento dos alunos e no seu futuro acadêmico e profissional. Garantir que se tenha um olhar crítico para os dados, procurando compreender o que podem representar num dado contexto e para um determinado indivíduo ou grupo de pessoas, seria seguir a ética enativista da intersubjetividade e da empatia.

Um exemplo prático seria o seguinte. Entre os recursos que se tornaram possíveis a partir de aprendizagem de máquina e *big data* estão os de rastreamento do comportamento de alunos quanto a vídeos de aulas numa plataforma online de aprendizagem. Pode-se saber se eles (e percentuais de grupos que) assistiram aos vídeos e quando, se pausaram e em quais momentos, se aceleraram os vídeos para assistir em menos tempo, se abandonaram as lições antes de terminar de assistir, entre outros dados. Como sugerem Mayer-Schönberger e Cukier (2014), identificar esses padrões nas atividades dos estudantes poderia, entre outros efeitos, ajudar professores de posse desses dados a ajustar suas aulas de acordo, fazê-los decidir reforçar conceitos que os alunos aparentemente não entenderam ou os levar a modular a forma de apresentar determinado tema. Seria uma orientação enativista fazer uso desses dados assim, como uma amostra, um dos fatores de orientação da aprendizagem, mas sem descartar a absoluta necessidade de que haja professores

para avaliar se realmente a aprendizagem sofreu um impacto positivo (ou mesmo qualquer tipo de impacto) a partir do uso de tais recursos. Não é possível, afinal, concluir que está de fato havendo aprendizagem ou que ela evoluiu de alguma maneira sem que se analise os dados dentro de um contexto e de um planejamento com vistas à aprendizagem. E sem que esses dados sejam analisados a partir da observação cotidiana e de uma troca dialógica com os alunos.

Sistemas artificiais são, em geral, valorizados por sua eficiência técnica. Isso significa que o sucesso desses sistemas depende da sua precisão, da sua estabilidade. A inconstância não costuma ser bem aceita nesse contexto: procura-se evitar erros, imprevistos e desvios, fatores indesejados quando o que se pretende é atender a determinadas demandas num curto espaço de tempo, absorvendo uma quantidade de dados vasta, atingindo-se cada vez mais velocidade, eficácia; a medida é a da produtividade. Sistemas são construídos para produzirem muito e serem estáveis, e isso os valoriza e os torna confiáveis. O que se pode dizer, por outro lado, sobre o sucesso humano? O sucesso humano em geral, e em particular o sucesso da aprendizagem formal, não raro é avaliado segundo parâmetros muito parecidos com aqueles que regem a eficácia de sistemas artificiais. Avaliações escolares são essencialmente balizadas por erros e acertos, por vezes com limitada margem para debates. Trabalhadores, por sua vez, são também quase sempre avaliados pelo tanto que produzem num determinado intervalo de tempo, num contexto em que o erro, em geral, aparece em oposição à produtividade, ao sucesso e à eficácia. É fácil medir esses erros e acertos de uma maneira que gera uma imensidão de dados sobre eles.

Por outro lado, a instabilidade, segundo o enativismo nos mostra, faz parte do funcionamento humano. Erramos o tempo todo porque somos imperfeitos, imprecisos e, o que é mais relevante, necessitamos da precariedade para agir no mundo. Precisamos, portanto, da instabilidade para nos mantermos autônomos. Diante dessa precariedade, a aprendizagem é contínua, nunca cessa; também se pode dizer que a aprendizagem é contínua justamente porque agimos na precariedade. Como propõe a filósofa enativista Stapleton (2021), porque a aprendizagem é contínua, ela gera transformações, mudanças; com isso, se, por um lado, é possível determinar fatos que se espera que alunos aprendam sobre história, por exemplo, não é tão simples desenhar resultados esperados quando se trata de

estimular o intelecto a partir do engajamento com disciplinas das humanidades como filosofia e literatura:

Esperamos para nossos alunos que eles nos deixem com mais do que a soma desses conhecimentos e seleção de habilidades. Esperamos que eles se tornem de alguma forma mais sábios, que cresçam como pessoas e como pensadores; que suas perspectivas sejam "transformadas". Tais resultados são difíceis de expressar em termos dos tipos de "resultados de aprendizagem" exigidos por instituições em suas descrições de cursos e programas porque não há um patamar "objetivo" em particular em que estamos tentando fazer com que as perspectivas dos alunos se transformem. Não estamos visando que eles internalizem e regurgitem nosso ponto de vista e nossa maneira de interagir com o mundo (pelo menos não completamente - há, é claro, aspectos de nossas maneiras de pensar que esperamos transmitir ao longo do curso de nosso ensino). Ao contrário, esperamos que eles desenvolvam suas próprias maneiras de pensar para serem mais inclusivos, mais discriminatórios e mais livres dos preconceitos particulares que trazem ao pensamento, seja em virtude do temperamento natural ou da educação. (Stapleton, 2021, p. 2)

Transformações na maneira de pensar e de agir são mais difíceis de medir e de prever, mas é possível observar quando elas acontecem. Na realidade, é sentindo que é possível percebê-las. Essa percepção depende da *experiência* de vivenciar as mais diversas situações. Stapleton (2021) – mencionando o trabalho de outra filósofa enativista que vem dedicando atenção às conexões entre enativismo e educação, Michele Maiese – chama atenção para o fato de que a literatura mais tradicional se refere a “tipos de aprendizagem”, mas que, nessa literatura, as mudanças de perspectiva implicadas na aprendizagem ficam apenas implícitas. O foco da autora, por outro lado, é justamente nas “mudanças fenomenológicas e epistemológicas de que a aprendizagem consiste” (Stapleton, 2021, p. 2).

Stapleton cita o fato de que experimentamos o mundo de maneiras diferentes quando somos crianças ou adolescentes. Ponderando as dificuldades de se medir os efeitos de uma mudança de perspectiva pela qual um indivíduo passe, a autora, porém, consegue ir mais longe na ilustração de um processo dessa natureza ao se referir a uma pesquisa feita com presidiários que fizeram um curso introdutório de filosofia. O método de ensino usado com eles baseou-se em perguntas que geravam debates sem que fosse perseguida uma conclusão comum para o grupo ou mesmo uma conclusão qualquer: o que era valorizado era a discussão em si. A seguir, descrevo o que ela relata em Stapleton (2021).

As regras eram simples: os participantes precisavam escolher um codinome, ou pseudônimo, que seria usado para identificá-los durante as sessões. Se quisessem falar, deveriam se dirigir ao professor e levantar a mão, e seriam

acolhidos ou não por ele dependendo do ponto onde o debate estivesse naquele momento. Se os alunos fossem chamados a falar, deveriam fazê-lo dentro do esquema “Concordo/Não concordo com x quando diz/dizem y porque z”. Havia, assim, uma limitação para os alunos, que precisavam expressar-se dentro desses parâmetros, bem como para o professor, que não podia ensinar de forma “tradicional”, como numa aula expositiva; não podia tampouco demonstrar qualquer julgamento ou interferir no curso da discussão, e precisava agir de modo a evitar selecionar perguntas “boas” ou que pudessem levar os alunos a conclusões que ele possivelmente desejaria que eles alcançassem. O professor teria que ter uma presença forte o suficiente para que os alunos entrassem no jogo e seguissem as regras, ao mesmo tempo que precisaria controlar os efeitos de sua personalidade ou carisma, atuando de uma forma mais “neutra” para não os conduzir pela mão. A presença dos alunos seria o fator mais importante.

A experiência é descrita por Stapleton (2021) como algo que resultou em uma mudança de perspectiva e também na postura dos alunos em situação de cárcere. É diferente, ela enfatiza, de apenas “adicionar algumas habilidades de pensamento crítico ao seu kit intelectual”. Os alunos foram instigados a pensar, por exemplo, sobre se deveriam aceitar tudo aquilo que ouvem, ou se deveriam questionar; e foram estimulados a não tomar como pessoais os diferentes pontos de vista que vinham dos colegas, por exemplo. Um dos resultados observados foi uma mudança de postura na hora de eles lidarem consigo mesmo e com outros, dentro e fora do espaço da prisão. Durante a prática, era estipulado que os alunos não deveriam falar da própria história, mas utilizar codinomes. Foi reportado pelo educador que essa prática ajudou os alunos a melhorar sua capacidade de argumentar e de falar por si próprios em entrevistas com agentes sociais ou com seus advogados. Esse aspecto da metodologia pode ser conectado à aprendizagem situada, mais precisamente ao conceito de *legitimate peripheral participation* (Lave, 1991). O conceito

(...) enfatiza a progressiva integração dos sujeitos nas comunidades da prática em que estes se movem, assim, do papel de observadores para o de agentes ou membros efectivos dessas comunidades. (Pessoa, 2002, p. 148)

Além disso,

[s]alienta-se também, no desenvolvimento de ambientes de aprendizagem situada, a possibilidade que se deve dar aos sujeitos de experimentar diversos papéis (Brown et al., 1989; Herrington & Oliver, 1997) como requisito necessário para uma realização com sucesso das várias tarefas a desempenhar. Esta possibilidade

acompanhada de uma reflexão sobre os diversos desempenhos é mesmo considerada por Brown et al. “one of the monumental task of education” (1989,39). (Pessoa, 2002, p. 148)

Stapleton (2020) indica que o caso da prisão é um exemplo de aprendizagem transformativa, tradução livre para *transformational learning* – “Uma mudança na perspectiva de alguém sobre o mundo que é, em algum sentido, profunda e fundamental”¹⁵⁸ (Stapleton, 2021, p. 8). Ressalta-se que a *teoria da aprendizagem transformativa* é uma teoria que se aplica a adultos, não a crianças, como coloca a autora. O motivo disso é que, na infância, a primeira perspectiva de significado ou sentido é formada, enquanto na vida adulta surge a possibilidade de transformação de perspectiva (Mezirow & Associados, 2000, p. xii, apud Stapleton, 2021, p. 9). No entanto, um aspecto que acredito que vale ser destacado aqui é que as ideias enativistas de *languageing*, *corporificação*, *experiência*, *emergência*, *autonomia* e outras servem não apenas para avaliarmos as transformações, em si, ocorridas nos alunos, mas para compreendermos como é que o professor, numa situação como a descrita, consegue perceber essas transformações nos seus alunos. Essa percepção, afinal, também é, em si, um processo de percepção fundamentado na experiência intersubjetiva corporificada, podendo ser estendida a toda situação de aprendizagem. Em outras palavras, tanto o que se aprende como a percepção daquilo que foi aprendido envolvem esses aspectos da cognição que o enativismo aprofunda. O professor, continuamente, aprende a observar a aprendizagem de seus alunos, e assim aprende como melhorar sua prática, também.

Cabe ao professor e a aquele que está se preparando para a docência perceber e avaliar a si mesmo e a aquilo que ocorre numa situação de aprendizagem. É útil, assim, que se conheça a importância de observar e sentir o desenrolar desses processos, apurando não apenas conhecimentos teóricos, mas a habilidade prática, mantendo o hábito de olhar atentamente para aquilo que acontece no contexto.

Pensar como professor é, como referimos, pensar de forma reflexiva num diálogo ou conversa com a prática, ou com diversas situações educativas construindo sentidos ou significados e, assim, conhecimento no sentido de uma acção competente. Pensar de forma reflexiva é ser capaz de exprimir afecto ou perplexidade perante situações imprevistas, é ser capaz de olhar as situações considerando perspectivas diversas, é ser capaz de perspectivar, situar ou enquadrar a situação complexa num problema tendo em conta, de forma não linear mas flexível, conhecimentos e experiências

¹⁵⁸ No original: *A change in one's perspective on the world that is in some sense deep and fundamental.* (Stapleton, 2020, p. 8).

passadas referenciados a um ideal ou ideia de ensino, aprendizagem e aluno. (Pessoa, 2002, p. 123-124)

No próximo tópico, irei explorar mais a questão acerca de como a intersubjetividade pode ser vista como uma aliada essencial quando se trata da inserção de tecnologias algorítmicas de machine learning na aprendizagem, com vistas a uma aprendizagem enativista.

6.3

Da interconexão entre mentes na busca de alternativas ao problema da autonomia

Na direção do olhar para a concepção de aprendizagem que emerge do enativismo, vale questionarmos um aspecto que ganha dimensão: como é que os sistemas de machine learning podem colaborar para que se tenha uma educação formadora e transformadora? A educação, afinal, envolve vários tipos de aprendizagem e de aquisição de conhecimento, mas está para além disso. Como claramente expõe Stapleton (2021),

Quando falamos de "educação" em vez de aprendizagem, treinamento ou desenvolvimento de habilidades, o que é que estamos tentando expressar? O que nos importa quando nos preocupamos com a educação das pessoas? Que mudanças queremos que elas experimentem? O conceito alemão de "Bildung" de alguma forma consegue captar isso um pouco melhor do que as palavras inglesas. (...) Esperamos para nossos alunos que eles nos deixem com mais do que a soma desses conhecimentos e uma variedade de habilidades. Esperamos que, de alguma forma, eles se tornem mais sábios, que cresçam como pessoas e como pensadores; que suas perspectivas sejam "transformadas" (Stapleton, 2021, p. 1)¹⁵⁹.

Se educar envolve uma série de processos de ensino e aprendizagem, talvez sistemas de machine learning possam ensinar algo a alguém, ou ajudar a ensinar algo. Será que podem? Considerando que temos um paradigma ainda forte na educação que é essencialmente cognitivista e preza pelo processamento de informações – mesmo quando se utiliza de saberes teóricos, metodologias e práticas construtivistas – existe no mínimo um trabalho com o qual sistemas que – por

¹⁵⁹ No original: *When we talk about “education” rather than learning, training, or skill development what is it that we are trying to express? What do we care about when we care about educating people? What changes do we want them to undergo? The German concept of “Bildung” somehow manages to capture this a little better than the English words. (...) We hope for our students that they will leave us with more than the sum of this knowledge and selection of skills. We hope that they somehow become wiser, that they grow as people and as thinkers; that their perspectives are “transformed”. (Stapleton, 2021, p. 1)*

exemplo – oferecem trilhas de aprendizagem podem colaborar: a transmissão de conteúdos. Esse é um tipo de aprendizagem que pode ser verificado com exames ou provas que geram dados; a partir desses dados, são oferecidos exercícios e atividades para esses estudantes aprimorarem tais conhecimentos. Nesses casos, vale ressaltar, se nada além disso for feito, essas máquinas serão empregadas para ensinar o mesmo conteúdo, de um mesmo currículo, que seria ensinado sem elas. Muda a tecnologia, mas não muda o paradigma; assim, não se altera a perspectiva ou o objetivo da aprendizagem. Nesse cenário, não importa tanto questionar a efetividade das máquinas; elas acabam, sim, por atender a aquilo para o qual foram designadas. Mas há mais nesse debate.

Sistemas de machine learning podem ajudar a ensinar, por vezes talvez possam ensinar algo “sozinhos” (quando os alunos usam esses sistemas para fazer atividades para absorver um conteúdo; mas isso precisa ser posto entre aspas especialmente porque humanos é que desenvolvem esses sistemas, então eles nunca são “sozinhos” de fato), mas é fato que sistemas artificiais não poderiam *educar* um ser humano. Educar é um atributo exclusivamente humano, o que é reforçado quando se parte de uma perspectiva enativista. Educar envolve mudanças de perspectiva, de tomadas de posição e direção – um movimento que coincide com o conceito enativista de *becoming*, ou tornar-se. O ser humano é inacabado, sempre está para se tornar, para se transformar. Por sua vez, a educação, como propõem Di Paolo, Cuffari e De Jaegher (2018, p. 309), não é uma opção, mas uma tarefa essencial para seres humanos, que implica em intervir no mundo¹⁶⁰.

Por conta desses aspectos – da educação como transformação, como intervenção no mundo e como algo intrinsecamente intersubjetivo – assume-se como foco aqui não as máquinas que ficam restritas ao espaço da sala de aula, mas os sistemas que nos acompanham no nosso cotidiano. Educação não é apenas o que se faz em sala de aula e as máquinas de ensinar, hoje, estão em todos os lugares da *algoritmosfera*. E se há algo na relação com tais máquinas que se distingue por completo da relação entre humanos, revelam-se aí aspectos tipicamente humanos que são caros à educação e também às aprendizagens que ela demanda. Em outras palavras, se aprender conteúdos não é suficiente, então os sistemas não poderiam

¹⁶⁰ “Education is not an option for human beings, but an essential duty, and to educate is to intervene in the world” (Di Paolo, Cuffari & De Jaegher, 2018, p. 309).

dar conta daquilo que a educação envolve. Surgem outras questões, para além daquela que expus no tópico anterior deste capítulo: qual o papel do professor e dos outros agentes envolvidos na educação, num contexto em que sistemas artificiais eventualmente poderiam ser usados para ensinar conteúdos? Quais outras aprendizagens uma educação que se pretende formadora e transformadora demanda, para além da aquisição de conteúdos?

Em meio a uma sociedade cada vez mais permeada por algoritmos, que geram um ambiente em que as trocas com o meio são desequilibradas e ameaçam a autonomia do cognoscente, o enativismo propõe uma perspectiva para a cognição humana que rompe com essa tendência. Ao propor a cognição calcada no corpo, uma forma de estar no mundo assentada no acoplamento com o ambiente, enativistas acentuam o papel da afetividade, da intersubjetividade e da experiência, que ganham um sentido alargado. E, assim, vão direto nos aspectos que são justamente como pontos cegos para um ambiente algorítmico. Para elaborarmos esse debate, precisamos, justamente, da intersubjetividade. Quando se trata de considerar as tecnologias digitais no contexto da aprendizagem, é preciso pensar: como elas podem contribuir para reforçar os laços intersubjetivos, em vez de afrouxá-los? E como o uso de sistemas artificiais pode estar alinhado a uma perspectiva que promove a intersubjetividade e coloca a interação humano-humano no centro? Este é um dos tópicos que será tratado aqui.

Como citado anteriormente, a interação é definida pelos enativistas como “a interdependência mútua (ou bidirecional, ou corregulada) dos comportamentos de dois agentes sociais” (Di Paolo, Cuffari & De Jaegher, 2018, p. 62). Como os autores explicam, os comportamentos específicos implicados nesse processo dependem da interação em si e da situação em que ela acontece. Existem vários graus ou níveis em que essa interação se desdobra, com resultados diferentes para a geração de sentido mútua envolvida em cada um desses níveis. Por exemplo, podemos dar uma orientação a uma pessoa que pediu ou ajudar alguém na rua; assim, estaremos influenciando na maneira como ela irá agir a partir da informação fornecida, e somente isto. Mas podemos avançar para graus mais elevados de *participatory sense-making*, isto é, graus em que o nível de participação é muito elevado. Isso acontece, por exemplo, quando um professor dá uma aula para uma turma de alunos ou quando escrevemos artigos científicos em parceria – este último

um exemplo dado pelos próprios enativistas, que são autores que desenvolvem muitas parcerias acadêmicas.

A interação entre agentes cognitivos, humanos, faz emergir algo que não existiria sem essa interação; algo que não pode ser inteiramente previsto, porque se molda a partir da interação e só existe justamente a partir dela e como resultado dela. É importante, então, não perder isso de vista quando se insere tecnologias digitais algorítmicas nos mais diversos contextos de aprendizagem. Por mais que se possa impactar negativamente a autonomia do cognoscente num ambiente algorítmico, uma vez que ele fica por vezes restrito a uma circularidade que não lhe trará tantas perspectivas emergentes e não lhe proporcionará tantas experiências quanto ele poderia ter engajando-se ou num ambiente não algorítmico ou ao lidar com outro cognoscente, pode-se sempre ter em mente que as trocas intersubjetivas são um recurso capaz de elevar essa autonomia novamente. Isto é, buscando um exemplo prático, se um aluno, por exemplo, é estimulado a fazer pesquisas num buscador online, e se depara com determinadas informações sobre o tema desejado, ele pode ser instigado pelo professor a procurar de outra maneira, a sair daqueles resultados que encontrou usando, para isso, outras palavras-chaves, por exemplo.

Pensando livremente sobre outros casos; um *chatbot* programado para propor e responder a perguntas sobre um idioma, como o inglês como língua estrangeira por exemplo, pode ser um recurso interessante de aprendizagem. Mas também pode ser bastante útil provocar nos alunos um debate sobre as próprias perguntas e respostas que observaram na interação com o sistema. Deve-se ter como princípio o fato de que somente da interação humana surgirão questões não programáveis (ou não programadas), capazes de expandir os horizontes e levar os alunos a domínios menos explorados. Usar o chatbot como recurso para aumentar o diálogo pode ser uma maneira de impulsionar a intersubjetividade, quebrar a circularidade, aumentar o grau de surpresa na situação e ainda despertar nos alunos a curiosidade de lidar com um sistema artificial ao mesmo tempo em que lhes são apresentadas as suas limitações. É importante, por exemplo, que sejam observadas as reações diante de recursos como bots e robôs humanóides, especialmente em crianças.

Como Colombetti (2014) indica, humanos atribuem expressividade mesmo a objetos inanimados, como bonecas, esculturas, troncos de árvores, nuvens e daí por diante, graças à nossa já mencionada *tendência empática* (Colombetti, 2014, p. 178). O episódio intitulado “Ashley Too”, da quinta temporada da série “Black

Mirror” (exibida no site de streaming *Netflix*), ilustra a possibilidade de uma adolescente apegar-se a um sistema de IA na forma de uma boneca, desenvolvendo afeição por ela. Apesar de trazer um caso ficcional, o episódio mostra como é importante cuidar para que os usuários de sistemas desse tipo não se esqueçam, e sejam informadas se necessário, de que estão lidando com algo capaz apenas de simular o comportamento humano (ver Selwyn, 2022).

6.3.1 Em busca de equilíbrio

Em 1998, Andy Clark e David Chalmers publicaram um artigo intitulado *The Extended Mind*. Nesse trabalho, os filósofos compartilharam com a comunidade científica a sua concepção de que a mente humana não se restringe ao cérebro, e nem ao corpo como um todo, mas inclui as tecnologias, sejam digitais ou não. Como Clark (2003) diria anos mais tarde, somos “ciborgues natos”. Isto é, temos condições de ampliar a nossa capacidade mental e de alcançar níveis cognitivos superiores quando nos acoplamos com elementos extraorgânicos. Se hoje isso pode parecer mais básico do que à época em que o artigo apareceu, provavelmente é porque já convivemos com smartphones há mais de 15 anos. Mas o artigo é anterior ao lançamento público dos primeiros celulares com conexão à internet, e antecede em quase dez anos o lançamento do primeiro iPhone.

A habilidade de Clark de antecipar tecnologias ou talvez a sua intuição o levou a escrever, no ano de 2001:

Ciborgues e *software agents* - Dois tipos de avanço tecnológico parecem prontos a estender a *mindfulness* humana de maneiras radicalmente novas. O primeiro, já familiar, mas que tem rapidamente crescido em ubiquidade e sofisticação, é exemplificado pelos assim chamados *software agents*. Um simples exemplo de um *software agent* seria um programa que monitora seus hábitos de leitura online, quais noticiários você frequentemente acessa e assim por diante, ou seus hábitos de compra online de CDs, e depois a busca por novos itens que combinam com seus atuais interesses. *Software agents* mais sofisticados podem monitorar leilões online, ofertas e vendas em seu nome, ou comprar e vender suas ações. (Clark, 2021, p. 92)

Aquilo que Clark chama de *software agent* poderia ser substituído por *plataforma baseada em algoritmos*. Ainda que hoje estejamos mais ouvindo música em sites de streaming do que adquirindo CDs, o que fazemos cada vez menos, trata-se de uma aposta muito certa: “programas” capazes de monitorar a nossa atividade online e até de tomar decisões por nós depois de “nos conhecer”. E Clark

continua, nos levando a pensar no que aconteceria com o passar do tempo em que experimentássemos uma evolução em conjunto com esses *software agents*:

Imagine que você começa a usar a internet aos quatro anos de idade. Software agents dedicados rastreiam e se adaptam aos seus interesses emergentes e às suas explorações aleatórias. Eles então ajudam a direcionar sua atenção a novas ideias, páginas na web e produtos. Pelos próximos 70 anos, você e seus software agents estão presos em uma dança complexa de mudança coevolutiva e aprendizagem, cada um influenciando e sendo influenciado pelo outro. Num caso assim, num sentido muito real, as entidades de software parecem mais uma parte de você do que uma parte do seu ambiente de resolução de problemas. O sistema inteligente que agora confronta o mundo amplo é o você-biológico-mais-os-software-agents. Esses pacotes externos de código estão contribuindo para aquele sistema inteligente um pouco como as várias funções cognitivas ativas em seu próprio cérebro. Eles estão constantemente trabalhando, contribuindo para o seu perfil psicologicamente emergente. (Clark, 2021, p. 93)

A visão de Clark tem um tom “otimista” com relação às tecnologias digitais e à circularidade gerada pela retroalimentação que mantemos com redes algorítmicas. Uma vez acopladas conosco, essas redes nos permitiriam ir além, facilitando as nossas vidas, nos permitindo acessar imediatamente aquilo de que precisamos e até antecipando coisas para nós. Faz parte da concepção da mente estendida essa compreensão de que o mundo ao redor de nós não apenas é útil para a nossa atividade cognitiva, mas pode ser moldado por nós de modo a nos ser útil. Podemos trabalhar o entorno para que esse espaço acompanhe as nossas necessidades. Organizamos uma mesa de trabalho de modo que os objetos de que precisamos estejam à mão para nós, por exemplo. Sem dúvida, essa é um aspecto positivo que pode emergir do uso que fazemos das tecnologias digitais; fica claro que há vantagens em ter acesso fácil a aquilo de que precisamos, e que pode agilizar os nossos esforços para cumprir tarefas.

Mas então por que foco nas preocupações com o crescimento de plataformas baseadas em algoritmos de machine learning, como as possibilidades de redução da autonomia do cognoscente? Por que, ou em qual sentido, essas tecnologias são diferentes de outras, que poderiam incrementar a nossa cognição talvez sem ressalvas?

Essencialmente, como procurei explicitar anteriormente, a minha preocupação vai no sentido de que as tecnologias baseadas em dados geram um ambiente em torno do ser cognoscente que ele, por vezes, não conhece e não sabe como funciona. Não é o cognoscente que está buscando organizar esse ambiente de uma determinada maneira, orientada às suas necessidades, como quando arruma

sua mesa de trabalho para ter objetos de escritório à mão, escolhe os sites que vai colocar nos favoritos para acessar mais rapidamente ou categoriza os arquivos de seu computador para fins parecidos com esses. Nessa economia de dados, quem – por assim dizer – “organiza os ambientes” são empresas detentoras da tecnologia, que armazenam dados de milhões de usuários para “devolver” a eles determinados conteúdos, baseados não apenas nos deles próprios, mas nos de outras pessoas. Como argumentei, isso afeta a autonomia do cognoscente essencialmente porque ele deveria conhecer as normas que regulam a suas trocas com o meio; mas, nesse caso, as normas são ocultas. Ele não sabe exatamente como são estipuladas, como são regidas.

Por vezes é conveniente, por exemplo, termos um buscador de internet que nos traga aquilo que, com base em interesses demonstrados em navegações anteriores, nos poderia ser útil ou prioritário. Mas as escolhas do cognoscente não são sempre facilitadas. Para ilustrar: os motores de busca procuram aquilo que está geograficamente perto de nós; por vezes isso pode ser útil, como quando queremos encontrar uma farmácia aberta. Mas o que acontece se quisermos pesquisar universidades em outro continente, por exemplo? A tendência é não termos tanto sucesso como teríamos se lá estivéssemos, presencialmente (por isso algumas pessoas têm preferido usar VPN). A questão é que esses mecanismos cada vez dependem menos de configurações a que temos acesso em nossos softwares e hardwares. Os algoritmos, por sua vez, são recortes que tosam aspectos importantes da realidade. Acabamos, portanto, à mercê de decisões das gigantes tecnológicas.

Esse argumento se relaciona com os pilares que Clark aponta como essenciais para que determinada tecnologia seja considerada uma extensão da mente humana. Na tentativa de criar parâmetros para isso, ele indica que é necessário que tal recurso tecnológico tenha sido endossado pelo próprio usuário, o que significa que se pode confiar nesse dispositivo. Um exemplo clássico é o caderno de Otto, que aparece no artigo de 1998 – Otto seria um paciente com Alzheimer que utiliza um caderno com suas próprias anotações para se guiar já que está perdendo a memória. O recurso deveria, ainda, ser acessível de maneira imediata, isto é, sem que o usuário sentisse uma interrupção entre o seu corpo e o recurso. Daí a classificação de bengalas para pessoas cegas, instrumentos musicais, próteses e celulares como extensões da mente humana – esses seriam recursos que se pode acessar automaticamente, de maneira fluida. Mas, como se nota, esses parâmetros não são

tão seguros quanto possa parecer; afinal, se uma tecnologia é ou não confiável e automática depende de uma série de fatores que, como mostram os sistemas artificiais movidos a redes neurais e algoritmos, podem variar.

No caso das tecnologias algorítmicas, um outro fator a ser considerado é que elas não são concebidas para atender individualmente a quem deseja potencializar suas capacidades cognitivas. Elas são desenhadas para operar com dados de milhões de usuários. Por mais que possam trazer algum benefício individualmente a cada um, são criadas para dar conta de algo que geralmente atende a interesses corporativos, empresariais ou governamentais. Os modelos matemáticos são desenhados a partir desses interesses. Esse é mais um ponto em que as tecnologias chamadas de “aprendizagem de máquina” se parecem com antigos recursos utilizados na educação, que citei no primeiro capítulo. Máquinas de escrever, por exemplo, eram propagandeadas como recursos potentes a partir de pesquisas contratadas para provar esse ponto. Seus usos poderiam ter algum benefício individual aos estudantes, mas atendiam, em última instância, a quem desejava vender mais máquinas.

Quanto maior a crença de que somos substituíveis por sistemas artificiais – crença esta que poderia se fortalecer caso se fechasse os olhos para as pesquisas sobre a cognição humana que ressaltam características como as que o enativismo aprofunda – tende a ser maior a permeabilidade da educação a invenções que muito prometem, mas pouco entregam; ou até entregam bastante, mas a um custo alto demais. Por vezes, para ter benefícios, fornecemos muitos detalhes de nossas vidas, personalidades, de nossa situação familiar e financeira, e o custo é a quebra de nossa privacidade, o uso de nossas informações de uma maneira irresponsável e antiética. Em vez de benefícios, pode-se ter dissabores muito graves. Uma pergunta pertinente é: vale a pena? Stockman e Nottingham (2022), citando Lyon (1994), indicam que, enquanto é promovido que tecnologias complexas oferecem conveniência, eficiência, segurança e uma incerteza cada vez menor, as pessoas acreditam nelas; mas também percebem, intuitivamente, que algo mudou e que pode haver algo errado: elas notam que é difícil definir o que está mudando, mas sentem que, seja o que for, pode afetar suas vidas. No caso da educação, algumas questões particulares devem ser consideradas e ponderadas no que diz respeito à vigilância proporcionada pela economia de dados em sistemas algorítmicos.

Dentro do contexto escolar, várias camadas de vigilância estão em funcionamento a todo momento. O aluno é geralmente sujeito à vigilância da escola, seja considerado como uma entidade de algum modo anônima (Dawson, 2006; Froelich, 2020), ou mais explicitamente na forma de equipe escolar individual, muitas vezes com um papel de ensino, segurança ou cuidado pastoral (Nagy, 2016). Por sua vez, os próprios professores e outros funcionários da escola também estão sujeitos a práticas de vigilância mais difundidas do que nunca (Página, 2017), assim como as escolas como um todo (Cullingford, 2016). Isso traz a impressão de todo o espaço educacional como uma objetividade calculável, e que deve ser vigiada de perto. (Stockman & Nottingham, 2022, p. 3)¹⁶¹

Ainda como enfatizam essas autoras, as tendências e práticas dentro da escola paralelizam com aquelas que estão do lado de fora, e que priorizam a previsibilidade e a proclamada eficiência técnica. Vastas bases de dados que traçam perfis pessoais pertencem a grandes corporações e departamentos governamentais, e mecanismos de captação de dados estão enraizados nos mecanismos por trás de mídias sociais, nos dispositivos que as pessoas usam para gravar a si mesmas e compartilhar esses conteúdos e nos recursos promovidos pela chamada internet das coisas, entre outros (Lyon, 1994, e Lyon, 2018, apud Stockman & Nottingham, 2022). As consequências dessa constante vigilância, como as autoras indicam, tendem a afetar os indivíduos em níveis diferentes, de acordo com características étnicas, sociais e ainda, por exemplo, a presença ou não de alguma deficiência – constituindo, assim, mecanismos potenciais de aumento da desigualdade e opressão. Por outro lado, as autoras (2022, p. 3) ponderam que recursos que promovem a vigilância são também aplicados como esforços para combater práticas de *cyberbullying*, monitorando as atividades de usuários de redes sociais ou, em outros casos, aumentar a segurança de crianças em escolas ou promover mecanismos capazes de checar como anda o bem-estar dos alunos.

Quanto ao empenho pedagógico, ele costuma ser justificado pelas mesmas razões que acompanhavam as antigas máquinas de ensinar. As autoras conectam esse suposto esforço a aquilo a que críticos têm se referido como “solucionismo” da indústria de EdTech, que oferece a si própria, com suas ideias e soluções, como

¹⁶¹ No original: *Within the school context, several layers of surveillance are in operation at any one time. The learner is commonly subject to surveillance by the school, either considered as a somewhat anonymous entity (Dawson, 2006; Froelich, 2020), or more explicitly in the form of individual school staff, often with a teaching, security or pastoral care role (Nagy, 2016). In turn, teachers and other school staff themselves are also subject to more pervasive surveillance practices than ever before (Page, 2017), as well as schools as a whole (Cullingford, 2016). It brings about the impression of the whole educational space as a calculative objectivity, and one which must be watched closely. (Stockman & Nottingham, 2022, p. 3)*

um caminho para resolver as lacunas que ela mesma afirma existirem na aprendizagem¹⁶². Certos recursos podem ser promovidos como gratuitos, mas ao final trazerem um custo (não financeiro) embutido bastante alto e ainda não precisamente estimado. Tendo em mente, assim, que é preciso cuidado com a pressa de assumir os efeitos positivos das tecnologias movidas a algoritmos, existem diversos esforços no sentido de criar regulações para que a economia de dados não seja tão desequilibrada, pendendo para o lado dos fornecedores e penalizando os consumidores, que são os fornecedores de dados. Em meio a esses esforços, alguns se dirigem a educar a sociedade quanto a esses riscos, outros se reúnem em torno da redação, análise e validação desses mecanismos reguladores, outros se dedicam à pesquisa, outros à comunicação das descobertas. Parecem ser empenhos válidos, uma vez que a única opção indiscutivelmente inviável é a aceitação acrítica e passiva dessa situação que nada tem de ingênua.

Voltadas especificamente às consequências da algoritmização para o campo da educação, Stockman e Nottingham (2022) exprimem preocupação no sentido de que currículos de alfabetização digital se restrinjam aos níveis instrumental e individual, correndo o risco de perder a perspectiva mais ampla da vigilância digital. Atentas aos pressupostos que sustentam a inteligência por trás de um currículo com essa finalidade, elas se referem a uma “ignorância epistêmica” descrita pelos autores Bhatt e Mackenzie (2019).

Os autores enfatizam a importância de um currículo que seja crítico e reflexivo, não somente sobre o ‘como’ da tecnologia digital, mas desenvolvendo uma conscientização maior sobre práticas sociais, forças econômicas e as assimetrias de poder relacionadas. (Stockman & Nottingham, 2022, p. 10)¹⁶³.

Para Selwyn (2022), pesquisador analítico das tecnologias educacionais, vários cuidados são importantes quando se trata de discussões sobre inteligência artificial na educação. Um deles é com as associações que costumam ser feitas por pessoas que trabalham com educação; elas tendem a presumir que formas de IA desenvolvidas para tarefas específicas são um passo significativo na direção da

¹⁶² Stockman e Nottingham (2002, p. 4) dizem: *This certainly corresponds to the EdTech industry’s ‘solutionism’ which therefore offers itself as a central need for either educational optimisation or enhanced safeguarding (Williamson & Hogan, 2020).*

¹⁶³ No original: *the authors emphasise the importance of a curriculum which is critical and reflective, not just on the ‘how to’ of digital technology, but in developing a greater awareness of social practices, economic forces, and the related asymmetries of power. (Stockman & Nottingham, 2022, p. 10)*

“general AI”, ou IA generalista. Esta é citada por Mitchell (2021) como uma das principais falácias envolvendo a IA. Viriam na esteira a crença na singularidade tecnológica, isto é, na superação humana pelos sistemas artificiais, e outros “sonhos de tecnologias capazes de alcançar níveis de senciência da inteligência humana” (Selwyn, 2022, p. 3). Mais importante e urgente, ele indica, é focar na IA que de fato já existe, com os “limites computacionais, materiais e metafísicos” inerentes às suas capacidades. Segundo o autor, é fundamental ter cuidado e até uma dose de suspeição quando se trata de tecnologias emergentes de IA na educação. Por mais que haja um sentido em ter atenção ao futuro que envolve tais tecnologias, por vezes, as discussões atuais focam em aspectos mais especulativos do que factuais, o que acaba distraindo de questões que já se fazem prioritárias.

6.4

Conclusão do capítulo

Sempre que alguém está empenhado no processo de aprender alguma coisa, essa experiência tem uma dimensão corporificada. Esta não é uma faceta da aprendizagem implicada apenas em atividades tipicamente físicas, portanto, mas em qualquer tipo de experiência de aprendizagem. O corpo é a interface com a qual se pode sentir, afetar-se pela experiência e ser impactado por ela. Isso acontece mesmo online e a distância. Há sempre algo na aprendizagem humana que é difícil de ser transposto para dados, simplesmente porque não é formalizável, não se encaixa nessa precisão; os melhores modelos matemáticos são incapazes de revelar aspectos essenciais da experiência de um corpo vivido, incluindo as suas expectativas, as suas sensações, as suas conexões com outras pessoas. Ainda que existam vertentes a defender que não é necessário nada biológico para que se tenha cognição, o argumento enativista indica que o elemento vital é essencial.

Outra dimensão fundamental de toda aprendizagem é a dimensão estética. “A cognição envolve uma sensibilidade estética que não pode ser codificada e elaborada em algoritmos” (Bannell, 2017, p. 32)¹⁶⁴. Quanto mais variadas as situações de aprendizagem que se possa experimentar, mais diversificadas serão também as habilidades recrutadas para esses contextos. Dewey (2008) falava em “qualidades sentidas”: são aquelas características de determinada situação que

¹⁶⁴ Ver também Leporace e Gondin, 2021.

somente podem ser percebidas na própria situação, até mesmo porque é a partir da interação que essas características se fazem presentes. A percepção que duas pessoas têm de um mesmo evento não é igual; ela depende do encontro entre o agente e aquele evento, do qual emerge a percepção. Mesmo uma pessoa que tenha duas ou mais vezes uma experiência semelhante poderá senti-la de maneiras muito distintas, já que o contexto influencia, bem como o que a pessoa aprendeu e sentiu com situações anteriores.

Um ponto nevrálgico desse pressuposto passa pelo fato de que seres humanos, enquanto cognoscentes, não se limitam a resolver problemas. A atividade mental humana, justamente por não se limitar a um processamento de informações, não se alimenta apenas de praticidade, redução de carga cognitiva e objetividade. Aprende-se pela experiência, e isso significa não apenas caminhar em linha reta, mas seguir (e criar) caminhos tortos, aparentemente “errados”, sinuosos. Improvisar, arriscar, ser capaz de mudar a rota também são atributos da cognição humana. Nem sempre a escolha que fazemos é pela rota mais fácil; podemos preferir variar um trajeto somente para respirar um “ar diferente” ou ver uma paisagem inusitada no caminho. Daí a importância da estética para a aprendizagem. A experiência cognitiva humana é, em si, uma experiência estética. Ser autônomo não é apenas ser capaz de resolver problemas objetivamente, mas vivenciar os sabores das descobertas, que, mesmo quando são amargos, são essenciais e caracterizam as nossas vivências. Conforme sinalizei no capítulo anterior, a circularidade da relação humana com sistemas de machine learning se torna negativa para a experiência quando a reduz.

Trazer uma perspectiva enativista para a aprendizagem abre uma oportunidade não de minimizar o tanto que as teorias de aprendizagem já evoluíram, muito menos negligenciar suas descobertas e avanços, mas de reorganizar ideias, compor com tradições e valorizar elementos que, apesar de importantes, poderiam receber mais atenção na educação. O intuito não é apenas gerar debate teórico, mas provocar mudanças práticas. Em reforço a isso, na busca de uma compreensão do possível papel de sistemas de machine learning quando se trata de aprendizagem humana, é preciso olhar justamente para aquilo que os algoritmos deixam de fora. É nesses espaços que estarão os aspectos da cognição humana que não podem ser delegados, que não admitem substituições. É com esse olhar que se pode, ainda, compreender as limitações dos sistemas artificiais, de uma maneira que contribua

para embasar decisões conscientes e seguras sobre como eles podem contribuir para os processos educacionais. Em reforço a isso, vale lembrar que a transição de um paradigma para outro leva tempo, e com isso convivem paradigmas diferentes em um mesmo período. Há diferentes percepções e motivações envolvidas. O enativismo mesmo mostra que esses movimentos não são lineares.

Não julgo ser impossível obter êxito, por exemplo, ensinando matemática ou português com o auxílio de plataformas adaptativas baseadas em machine learning. Porém, mesmo que se consiga bons resultados, não se pode esquecer que estaremos assim ensinando matemática e português da mesma forma já ensinada nas escolas, seguindo mais ou menos os mesmos métodos de avaliação, ainda que de maneira automatizada. Também estaremos operando dentro de uma determinada lógica sobre aquilo que significa ter sucesso e obter resultados. É importante não perdermos a noção disso, uma vez que permanece a questão da mudança de paradigma e a demanda por uma forma de aprender e de ensinar que considere a cognição para além do cérebro. Ainda assim, como já colocado, mudanças de paradigma não são rápidas. Enquanto vivemos a transição, é esperado que muitas crenças e práticas convivam.

Também não penso, de maneira alguma, que não exista por exemplo uma praticidade proporcionada por termos tecnologias algorítmicas trabalhando em acoplamento conosco. Mas devemos ter uma noção alargada daquilo que está acontecendo para podermos ser críticos e cuidadosos; devemos ser educados para esse contexto digital, para esse ambiente algorítmico movido a dados que povoamos e que ajudamos a criar; precisamos trabalhar para que a identidade das pessoas não seja reduzida a aquilo que os algoritmos podem encapsular; precisamos garantir que professores compreendam o seu papel em um contexto epistemológico no qual ferramentas “ensinam” e surpreendem pelo “tanto” que sabem, mas apenas aparentemente; precisamos que se compreenda o que significa ter seus dados disponibilizados nas redes, as questões da privacidade, da tomada de decisão, da ética; precisamos que as pessoas saibam como acessar os caminhos alternativos, mais longos, menos “diretos” e que não estejam dados, mas que sejam justamente aqueles caminhos que podem abrir outros. Isto é, precisamos compreender a circularidade que mantemos com a *algoritmosfera* para que ela possa resultar em algo que de fato signifique crescimento cognitivo, sem correr o risco de ser apenas rapidez ou eficiência para resolver problemas. A *algoritmosfera* depende de nós

para se manter e podemos e devemos, assim, moldá-la em nosso favor. A noção de educação que emerge da perspectiva enativista e as ideias de *participatory sense-making* e de *sense-making participativo* apoiam essa perspectiva.

7

Considerações finais

No decorrer deste trabalho, procurei demonstrar como é oportuno analisar o significado e os limites da *aprendizagem* em sistemas artificiais quando se deseja compreender a *aprendizagem humana* e as suas possibilidades. É também importante conhecer os fatores que diferenciam essas esferas para que se compreenda como elas podem se articular, incluindo-se aí uma compreensão sobre como machine learning pode estar presente na aprendizagem humana. Conforme procurei explorar, o enativismo sugere uma *conceituação* da cognição que envolve um fazer sentido do mundo situado, corporificado e afetivo. A aprendizagem de máquina distancia-se do *sense-making* enativista porque essa perspectiva se traduz, essencialmente, em *corpo orgânico em movimento*, articulado a outros corpos orgânicos em movimento.

Quando compreendemos que sistemas artificiais não podem ser considerados agentes cognitivos ou *sense-makers* e o porquê disso – eles não têm experiência sob a perspectiva enativista do conceito, nem autonomia, nem outros atributos advindos de um corpo acoplado com o mundo – pode-se concluir que esses sistemas não aprendem como humanos aprendem. O significado de aprendizagem em machine learning não deve, portanto, ser confundido com o significado de aprendizagem em humanos. O contrário, porém, ainda demanda mais cuidado: seres humanos não aprendem como máquinas, mas o cognitivismo, concepção ainda preponderante, preza pelo processamento de informações, a inferência e as representações mentais. E nós estamos, como o enativismo mostra, muito além disso.

A partir do estudo da perspectiva filosófica para a cognição definida como enativismo, pode-se dizer que o indivíduo *aprende* a partir de seu acoplamento com o mundo. Essa aprendizagem distancia-se muito do sentido de aprendizagem inerente aos sistemas de ML – que não podem ser considerados agentes, já que são sistemas sem *experiência* no sentido enativista e desconectados da *experiência* de outros agentes; além disso, são orientados pelas escolhas daqueles que são responsáveis por criar esses sistemas, os quais deliberam parâmetros para eles por meio de algoritmos e modelos. Devemos ainda nos lembrar, mais uma vez, que algoritmos não são neutros (O’Neil, 2016); por definição, são fragmentos do mundo, sempre parciais, e por isso deixam de fora uma parte grande do cenário

completo. O problema é que, por vezes, o ponto cego de um sistema baseado em algoritmos corre o risco de ser justamente a parte mais importante do contexto, que não poderia ser deixada de lado. Nesse sentido, o mundo ainda é o melhor modelo de si mesmo, como diria Brooks (1991). Para sustentar sua autonomia, não apenas buscando sobrevivência, mas também tentando encontrar as melhores respostas para as nossas expectativas, como cognoscentes devemos ter o mundo a nosso alcance; mas não um mundo reduzido, que deixa de fora partes essenciais de nossas mentes – nossos corpos e suas dimensões orgânica, sensório-motora e intersubjetiva e todas as possibilidades de engajamento com o mundo e com outros que emergem daí.

Estamos imersos em redes corporais, materiais, em constante enlace, mantidas por conexões entre sujeitos cognoscentes. Se circunscritos a um conceito de mente que se restringe à cabeça – ao qual o enativismo é uma alternativa teórico-prática – essencialmente reduzimos os processos de aprendizagem à atividade cerebral de processamento de informações e a um sistema interno e fechado em um sujeito. Equiparar a atividade de dar sentido ao mundo à mera decodificação e ao processamento de informações corresponde a uma classificação daquilo que está presente no mundo que se encaixaria em uma ontologia cartesiana atomizada, onde as partes são percebidas antes do todo. Nós, entretanto, lidamos com o mundo como ele é, como um todo, um *continuum*, com suas instabilidades e imprecisões, como colocaria Dewey (2008). E é justamente a nossa experiência que nos permite perceber o ambiente a partir de sua riqueza com dimensões multifacetadas, o que nos possibilita até mesmo a proceder a essa análise das partes se assim desejarmos.

É por meio desse processo que podemos desenvolver nossas capacidades como *sense-makers* autônomos. A célula que “segue” partículas de açúcar para sua sobrevivência não precisa saber a fórmula do açúcar para se orientar a uma de suas moléculas; ela simplesmente faz esse movimento. De maneira similar, nós não

precisamos saber qual é a composição química de algo para lidar com essa substância em nosso cotidiano (ainda que pudéssemos saber, se fôssemos cientistas em um laboratório por exemplo). Para esses constantes contatos e acoplamentos, contamos com nossas dimensões corporais, nossas habilidades de perceber e daí por diante. Para fazer sentido do mundo, e para agir enquanto fazemos sentido do mundo, devemos ser capazes de percebê-lo como um todo, não a partir de uma perspectiva pré-determinada externa a nós, como um sistema que não tem nenhuma

participação sobre a maneira como se organiza – e, por isso, não pode ser considerado autônomo. Nós construímos o caminho ao caminhar, como nos lembram Varela, Thompson e Rosch (2016).

No caso de machine learning, se não conhecermos cada sistema individualmente, ao menos podemos estar conscientes acerca da lógica geral desses sistemas; eventualmente, sabendo os princípios que os guiam, é possível contrastá-los com a nossa própria atividade cognitiva. Isso nos ajuda a não perder de vista as nossas habilidades e características de *sense-makers autônomos*. Se for possível argumentar, como procurei fazer, que a nossa *autonomia* pode ser ameaçada quando lidamos com sistemas de ML, acredito que ter essa consciência pode ao menos nos preparar para isso. O espaço para que emergja o novo, a surpresa, o inesperado, e o espaço para as experiências corporais, intersubjetivas, é essencial para a educação e para a formação do ser humano. Ao experimentar o mundo, o cognoscente exerce a sua autonomia, buscando as melhores possibilidades de acoplamento com o meio. Os ambientes criados após o advento dos sistemas de ML não mantêm as mesmas características dos ambientes criados a partir do *sense-making participativo* de sujeitos com corpos físicos interconectados. Talvez esse seja o mais importante aspecto, ou ao menos um dos mais importantes, para uma análise focada nas consequências possíveis do uso desses sistemas para fins de aprendizagem e desenvolvimento. Parece que a melhor extensão para a mente de um ser humano¹⁶⁵, mais do que a tecnologia, é *o outro*¹⁶⁶. Isso tem implicações quando se pensa sistemas de ML como plataformas de aprendizagem que, se por um lado são impulsionadas pelo discurso de atender aos alunos de forma individual, como as antigas máquinas de ensinar, por outro podem acabar massificando e aumentando a distância entre alunos e professores.

Autonomia, no sentido enativista, existe e é mantida no acoplamento com o ambiente que, apesar de instável e sujeito a mudanças, não é hostil à *experiência*. Esse acoplamento, afinal de contas, fornece os meios para um constante tornar-se, o *becoming* que o enativismo aborda (Di Paolo, 2020). Como seres vivos, especialmente seres linguísticos, nós no encontramos em um fluxo incessante. Esse

¹⁶⁵ Faço aqui referência ao trabalho de Andy Clark, proponente da mente estendida; segundo ele, somos ciborgues natos devido a nossa capacidade de, desde sempre, nos acoplar a tecnologias de uma maneira que estende as nossas capacidades mentais. Ver Clark (2003; 2014).

¹⁶⁶ Poderia ser mencionado aqui o conceito de empatia tal como desenvolvido por fenomenologistas (Zahavi, 2021).

estado de *tornar-se/becoming* demanda a busca constante pelo melhor acoplamento entre o sujeito e o mundo. Então, se machine learning, hoje, em princípio oferece possibilidades restritas para o sujeito acoplar-se com o mundo, isso deve ser levado em consideração ao serem tomadas decisões envolvendo a inclusão desses sistemas em experiências educacionais, que têm como objetivo último contribuir para a formação humana. Esse aspecto não pode ser desassociado de implicações éticas e possíveis ameaças à integridade e à privacidade humana. Sistemas de machine learning têm surgido no contexto educacional como tecnologias de ensino e aprendizagem, além de estarem sendo usados para avaliações de alunos e professores; em todos esses casos, essas aplicações têm grande potencial de afetar o futuro dessas pessoas.

Coeckelbergh (2020) alerta para a vulnerabilidade a que usuários de sistemas de ML têm sido expostos, tendo em conta que nossos dados são coletados de uma maneira que nem sempre preserva a nossa privacidade e o nosso direito de saber o que acontece com essas informações. Deveríamos ter e a possibilidade de fazermos escolhas de acordo. Poderíamos reagir negativamente a isto se soubéssemos o que está acontecendo, e então não consentir, por exemplo. Quando Coeckelbergh (2020, p. 111) diz que “Machines can be agents but not moral agents since they lack consciousness, free will, emotions, the capability to form intentions, and the like”, e comenta como a ciência de dados recruta humanos a preparar, analisar e interpretar os resultados apresentados por redes de ML, ele está indicando que máquinas não devem ser deixadas sozinhas em processos de decisão que podem impactar vidas humanas. Ainda que sistemas de ML terminem por “tomar decisões”, num sentido meramente estatístico, faltam-lhes recursos que são essenciais para esses processos de tomada de decisão. Talvez se possa argumentar que a humanos falte também senso de justiça ao deliberar em processos semelhantes, mas ao menos podemos recorrer à nossa capacidade de *pensar sobre o pensar* e questionar-nos acerca do que seria uma sociedade justa. Na contramão disso, sistemas de ML seguem sendo caixas pretas, ou ao menos caixas cujos conteúdos e resultados não podem ser controlados por nós.

Além disso, e de forma primordial, sob o ponto de vista enativista – encapsulado no conceito de *fazer sentido de maneira participativa* – quando lidamos com humanos, podemos pelo menos ajustar as expectativas e ter um diálogo de humano para humano, de mente para mente, de corpo para corpo. Isto

só acontece quando se lida com outros; não ocorre quando se lida com sistemas artificiais cegos, sem mente, descorporificados e sem a capacidade de sentir. O problema é que estes sistemas podem levantar paredes de concreto ao nosso redor, em vez de abrir estradas fluidas para que continuemos a construir nossos próprios caminhos; neste sentido, nossos dados tornam-se *another brick in the wall*, como diria a música do Pink Floyd, ou apenas *mais um tijolo numa parede*. Talvez devêssemos nos preocupar em olhar entre as falhas nos tijolos.

8

Referências Bibliográficas

- APA / American Psychological Association. Self-instructional materials and devices. **American Psychologist**, 16(8), 512-513, 1961. <https://doi.org/10.1037/h0043852>.
- BANNELL, R. I.; MIZRAHI, M.; FERREIRA, G. (Orgs.) **(Des)educando a educação: Mentas, Materialidades e Metáforas**. Rio de Janeiro: PUC-Rio, 2021.
- BANNELL, R. I. *et al.* **Educação no Século XXI: cognição, tecnologias e aprendizagens**. Rio de Janeiro: Editora Vozes, 2016.
- BANNELL, R. I. Linguagem, articulação e educação. In: **Hermenêutica e Educação**. Batista, G. S., Nascimento, C. R. e Pimenta, A. R. (Orgs.), 2022.
- BANNELL, R. I. Uma faca de dois gumes/A double-edge sword. In: FERREIRA, Giselle M. S.; ROSADO, Luiz A. S.; CARVALHO, Jaciara S. (Org.) **Educação e Tecnologia: abordagens críticas**. Rio de Janeiro: SESES – Sociedade de Ensino Superior Estácio de Sá, 2017, p. 17-82.
- BANNELL, R. I. Out of our Minds? Learning beyond the brain. In Leporace, C., Bannell, R.I., Rodrigues, E., Santos, E. (Orgs.) **A Mente Humana para além do Cérebro: Perspectivas a partir dos 4Es da cognição**. Coimbra: Portugal: IPCDHS, 2019.
- BENJAMIN, L. T. A history of teaching machines. **American Psychologist**, 43(9), 703–712, 1988. <https://doi.org/10.1037/0003-066X.43.9.703>
- BIESTA, G. Against learning. Reclaiming a language for education in an age of learning. **Nordisk Pedagogik**, Vol. 25, pp. 54–66, 2005. Oslo. ISSN 0901-8050.
- BOALER, J. Seeing is Achieving: the Importance of Fingers, Touch, and Visual Thinking to Mathematics Learners. In: Macrine, S. L. e Fugate J. M. B. (Edit.) **Movement Matters – How Embodied Cognition Informs Teaching and Learning**, pp. 121-130.
- BROOKS, R. Intelligence without reason. In: **Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI-91)**, pp 569–595, Sydney, Australia: 1991.
- BROOKS, R. 1990. Elephants don't play chess. In: P, Maes (ed.) **Designing Autonomous Agents**, pp 3–15, MIT Press
- CANDAU, V. M. **Ensino Programado – Uma nova tecnologia didática**. Rio de Janeiro: Inter edições, 1969.
- CANGELOSI, A.; SCHLESINGER, M. **Developmental Robotics: from Babies to Robots**. Cambridge: MIT Press, 2015.
- CHALMERS, A. F. **O que é ciência afinal?** Tradução de Raul Filker. Brasília: Editora Brasiliense, 1993.
- CLARK, A. **Being there: putting brain, body, and world together again**. Cambridge, MA: MIT Press, 1998.
- CLARK, A. **Natural-Born Cyborgs. Minds, Technologies and the Future of Human Intelligence**. New York: Oxford University Press, 2003.

CLARK, A. **Supersizing the Mind – Embodiment, action, and cognitive extension** Oxford: Oxford University Press, 2011.

CLARK, A. **Mindware**. Oxford: Oxford University Press, 2014.

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

CLARK, A. Tecnologia cognitiva – Para além do cérebro nu. In: BANNELL, R. I.; MIZRAHI, M.; FERREIRA, G. (Orgs.) **(Des)educando a educação: Mentres, Materialidades e Metáforas**. Rio de Janeiro: PUC-Rio, 2021. Tradução de Camila De Paoli Leporace.

COECKELBERGH, M. **AI Ethics**, Cambridge, Massachussets, London, England, The MIT Press, 2020.

COLOMBETTI, G. **The Feeling Body – Affective Science Meets the Enactive Mind**. Cambridge: MIT Press, 2014.

DAMÁSIO, A. R. **O erro de Descartes**. São Paulo: Companhia Das Letras, 2012.

DAMÁSIO, A. R. **The Strange Order of Things – Life, Feeling and the Making of Cultures**. New York: Pantheon Books, 2018.

DE JAEGER, H. Social understanding through direct perception? Yes, by interacting. **Consciousness and Cognition**. Vol. 18, ed. 2, p. 535-542, 2009. <https://doi.org/10.1016/j.concog.2008.10.007>

DE JAEGER, H. How we Affect Each Other. **Journal of Consciousness Studies** 22 (1-2) p. 112-132, 2015. <https://doi.org/0803973233>

DE JAEGER, H.; DI PAOLO, E. A. Participatory Sense-making: An enactive approach to social cognition. **Phenomenology and Cognitive the Sciences**, 6, pp. 485-507, 2007.

DE JAEGER, H. Making Sense in Participation: An Enactive Approach to Social Cognition. In: **Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions**. F. Morganti, A. Carassa, G. Riva (Eds.) Amsterdam, IOS Press, pp. 33-47, 2008.

DEEP MIND. **Making History** – AlfaGo is the first computer program to defeat a professional human Go player, the first to defeat a Go world champion, and is arguably the strongest Go player in history. Disponível em <https://deepmind.com/research/case-studies/alphago-the-story-so-far>. Acesso em 1 de setembro de 2021.

DELLERMANN, D.; EBEL, P.; SÖLLNER, M.; LEIMEISTER, J. M. (2019). Hybrid Intelligence. **Business and Information Systems Engineering**, 61(5), 637–643. <https://doi.org/10.1007/s12599-019-00595-2>

DESCARTES, R. **Regras para a orientação do espírito**. São Paulo: Martins Fontes, 2012.

DEWEY, J. **Experience and Nature - The Later Works of John Dewey, Volume 1, 1925 - 1953: 1925**. 1ª ed. Illinois: Southern Illinois University Press, 2008.

DI PAOLO, E. Autopoiesis, adaptivity, teleology, agency. **Phenomenology and the Cognitive Sciences**. Sci. 4, 429–452, 2005. doi: 10.1007/s11097-005-9002-y

DI PAOLO, E. Extended Life. **Topoi** 28, 9, 2009. <https://doi.org/10.1007/s11245-008-9042-3>

- DI PAOLO, E. Prefácio. In: ROHDE, M. **Enaction, Embodiment, Evolutionary Robotics**. Paris: Atlantis Press, 2010.
- DI PAOLO, E. A Conceção Enativa da Vida. In: BANNELL, R. I.; MIZRAHI, M.; FERREIRA, G. (Orgs.) **(Des)educando a educação: Mentres, Materialidades e Metáforas**. Tradução de Camila De Paoli Leporace. Rio de Janeiro: PUC-Rio, 2021.
- DI PAOLO, E. Enactive becoming. **Phenomenology and the Cognitive Sciences**, 1–27, 2021.b <https://doi.org/10.1007/s11097-019-09654-1>
- DI PAOLO, E.; ROHDE, M.; DE JAEGHER, H. Horizons for the Enactive Mind: Values, Social Interaction, and Play. In: **Enaction – Towards a New Paradigm in Cognitive Science**. STEWART, J.; GAPENNE, O; DI PAOLO, E. London: The MIT Press, 2010.
- DI PAOLO, E; BUHRMANN, T.; BARANDIARAN, X. **Sensorimotor Life – An Enactive Proposal**. Oxford: Oxford University Press, 2017.
- DI PAOLO, E. A.; CUFFARI, E. C.; DE JAEGHER, H. **Linguistic Bodies. The Continuity between Life and Language**. Cambridge: MIT Press, 2018.
- DI PAOLO, E.; THOMPSON, E. The enactive approach. In: SHAPIRO, L. **The Routledge Handbook of Embodied Cognition**. London: Routledge Taylor & Francis Group, pp. 86-96, 2014.
- DOMINGOS, P. **The Master Algorithm. How the Quest for the Ultimate Learning Machine Will Remake Our World**. New York: Basic Books, 2015.
- DREYFUS, H. L; DREYFUS, S. E. **Mind over machine: The Power of Human Intuition and Expertise in the Era of the Computer**. New York: The Free Press, 1986. ISBN 0-7432-0551-0
- DREYFUS, H. L.; DREYFUS, S. E. Peripheral Vision: Expertise in Real World Contexts. **Organization Studies**, London: SAGE Publications, v. 26, n. 5, p. 779–792, 1 maio 2005. DOI 0.1080/09515080701239510. Disponível em: <<https://journals.sagepub.com/doi/10.1177/0170840605053102#articleCitationDownloadContainer>>. Acesso em 20 de agosto de 2020.
- DREYFUS, H. L. Overcoming the Myth of the Mental: How Philosophers Can Profit from the Phenomenology of Everyday Expertise. **Proceedings and Addresses of the American Philosophical Association**, [S. l.], v. 79, n. 2, 1 nov. 2005. American Philosophical Association, p. 47-65. Disponível em: <<http://www.jstor.org/stable/30046213>>. Acesso em 10 de dezembro de 2020.
- DREYFUS, H. L. Refocusing the question: Can there be skillful coping without propositional representations or brain representations? **Phenomenology and the Cognitive Sciences**, Netherlands: Kluwer Academic Publishers, v. 1, p. 413-425, 2020.
- DREYFUS, H. L. Why Computers Must Have Bodies in Order to Be Intelligent. **The Review of Metaphysics**, [S. l.], v. 21, n. 1, ago. 1967. p. 13-32. Disponível em: <<http://www.jstor.org/stable/20124494>>. Acesso em 10 de dezembro de 2020.
- DREYFUS, H. L. Why Heideggerian AI Failed and How Fixing it Would Require Making it More Heideggerian. **Philosophical Psychology**, [S. l.], v. 20, n. 2, p. 247-268, 23 abr. 2007. DOI 0.1080/09515080701239510. Disponível em: <<http://dx.doi.org/10.1080/09515080701239510>>.

DREYFUS, H. L. **What Computers Still Can't Do – A Critique of Artificial Reason**. Cambridge, MA: MIT Press, 1992.

DREYFUS, H. L. **O que os computadores não podem fazer – Uma crítica da Razão Artificial**. Rio de Janeiro: A Casa do Livro Eldorado, 1975.

DREYFUS, H. L. **On the Internet**. Londres e Nova York: Routledge, 2009.

DREYFUS, H. L. **Skillful Coping: Essays on the Phenomenology of Everyday Perception and Action**. Oxford: Oxford University Press, 2014.

FERREIRA, CARVALHO E LEMGRUBER. Heidegger, Tecnologia e Arte: impasses contemporâneos. **Revista Enunciação**, v. 3, p. 26-39, 2018. [HTTP://DX.DOI.ORG/ 10.5935/2238-1279.20190001](http://dx.doi.org/10.5935/2238-1279.20190001)

FORBUS, K. D.; FELTOVICH, P. J. **Smart Machines in Education**. Cambridge, MA: MIT Press, 1997.

GALLAGHER, S. A.; BOWER, M. (2014). Making enactivism even more embodied. **Avant**, 5 (2), 232-247.

GIBSON, J. J. **The Ecological Approach to Visual Perception**. New York: Psychology Press, 1986.

HOLMES, W.; BIALIK, M.; FADEL, C. **Artificial Intelligence in Education – Promises and Implications for Teaching and Learning**. Boston: The Center for Curriculum Redesign, 2019.

HOW KISMET WORKS. Disponível em <https://www.youtube.com/watch?v=KwgOmJwzuc>. Acesso em 10 de novembro de 2022.

HUMAN BRAIN PROJECT. Disponível em <https://www.humanbrainproject.eu/en/>. Acesso em 20 de dezembro de 2021.

HURLEY, S. Perception and Action: Alternative Views. **Synthese** 129, 3–40, 2001. <https://doi.org/10.1023/A:1012643006930>

INTERNET 2020 - Cresce a liderança dos smartphones. **Marktest**, 8 de setembro de 2020. Disponível em <https://www.marktest.com/wap/a/n/id~269e.aspx>. Acesso em 10 de outubro de 2022.

JOHNSON, M. **The meaning of the body: aesthetics of human understanding**. Chicago & London: The University of Chicago Press, 2007.

JOHNSON, M. **Embodied mind, meaning and reason - how our bodies give rise to understanding**. Chicago: University of Chicago Press, 2017.

JOHNSON, M. **The aesthetics of meaning and thought – the bodily roots of Philosophy, Science, Morality and Art**. Chicago: University of Chicago Press, 2018.

JOHNSON, M. A corporificação da linguagem. In BANNELL, R.I., MIZRAHI, M. & FERREIRA, G. (Orgs). In: **Deseducando a Educação: mentes, Materialidades e Metáforas**. Editora PUC-Rio, 2021. Tradução de Camila De Paoli Leporace.

JOHNSON-GLENBERG, M. Evaluating Embodied Immersive STEM VR Using the Quality of Education in Virtual Reality Rubric (QUIVRR). In: Macrine, S. L. e Fugate J. M. B. (Edit.) **Movement Matters – How Embodied Cognition Informs Teaching and Learning**, pp. 237-260.

JOLDERSMA, C. Neuroscience, Education, and a Radical Embodiment Model of Mind and Cognition. In C. Mayo (Ed.), **Philosophy of education**, 2013.

JONAS, H. **The phenomenon of life: Towards a philosophical biology**. Evanston, IL: Northwestern University Press, 1966.

KELLY, S. D. Grasping at Straws: Motor Intentionality and the Cognitive Science of Skilled Behavior. In: WRATHALL, M.; MALPAS, J. (ed.). **Heidegger, Coping, and Cognitive Science: Essays in Honor of Hubert L. Dreyfus**. 2. ed. Cambridge, MA: MIT Press, 2000. cap. 8, p. 161-177.

KNOX, J.; WILLIAMSON, B.; BAYNE, S. Machine behaviourism: future visions of 'learnification' and 'datafication' across humans and digital technologies. **Learning, Media and Technology**, 45:1, 31-45, 2020. DOI: [10.1080/17439884.2019.1623251](https://doi.org/10.1080/17439884.2019.1623251)

KURZWEIL, R. **The Singularity Is Near: When Humans Transcend Biology**. Londres: Duckworth Overlook, 2008.

LEPORACE, C. P. Espelhos do mundo? Uma perspectiva da percepção humana a partir das ideias da mente enativa e da mente estendida. In: LEPORACE, C. P.; BANNELL, R. I.; RODRIGUES, E.; SANTOS, E. J. R. S. (Org.). **A Mente Humana para Além do Cérebro - Perspectivas a partir dos 4Es da Cognição**. 1ed.Coimbra: Universidade de Coimbra, 2019, v.1, p. 37-59.

LEPORACE, C. P. **Somos todos ciborgues: a tese da mente estendida e as tecnologias digitais na educação**. Dissertação (mestrado). 162p. Departamento de Educação da Pontifícia Universidade Católica do Rio de Janeiro, 2019.

LEPORACE, C. P. Another Brick in the Wall: Threats to our autonomy as sense-makers when dealing with machine learning systems. **PERSPECTIVA FILOSOFICA** [S.l.], v. 49, n. 5, p. 200-230, dez. 2022 (a). ISSN 2357-9986.. 2022 (a) <https://doi.org/10.51359/2357-9986.2022.252618>.

LEPORACE C. P. O que os computadores continuam não conseguindo fazer, 50 anos depois: a aprendizagem sob a perspectiva da fenomenologia do cotidiano de Hubert Dreyfus. **REVISTA PORTUGUESA DE EDUCAÇÃO**, 35(2), pp. 05–23, dez. 2022(b). <https://doi.org/10.21814/rpe.23877>

LEPORACE, C. P.; BANNELL, R. I. (Org.); RODRIGUES, E. M. S. (Org.); SANTOS, E. J. R. (Org.). **A Mente Humana para Além do Cérebro - Perspectivas a Partir dos 4Es da Cognição**. 1. ed. Coimbra: Universidade de Coimbra, 2019. v.1. 174p

LEPORACE, C.; GONDIN, V. C. Experimentar o mundo a partir do corpo: estética como uma dimensão da cognição humana. **Educação On-Line**, Rio de Janeiro, v. 16, n. 38, set-dez 2021, pp. 227-244.

LYNCH, M. The Benefits and Limitations of Machine Learning in Education. **The Tech Advocate**, 17 de maio de 2018. Disponível em <https://www.thetechadvocate.org/the-benefits-and-limitations-of-machine-learning-in-education/>. Acesso em 10 de setembro de 2022.

LYON, D. **The Electronic Eye: The Rise of the Surveillance Society**. Minneapolis: Polity Press/Blackwell, 1994.

MACHINE LEARNING IN EDUCATION. Amazon AWS. Disponível em <https://aws.amazon.com/pt/education/ml-in-education/>. Acesso em 10 de julho de 2021.

MACHINE LEARNING IN EDUCATION - : Benefits and Opportunities to Explore. Intellias – Global Technology Partner. Disponível em <https://intellias.com/benefits-of-machine-learning-in-education/>. Acesso em 10 de julho de 2021.

MACKENZIE, A. The production of prediction: What does machine learning want? **European Journal of Cultural Studies**. 2015;18(4-5):429-445. doi:10.1177/1367549415577384.

MATURANA, H; VARELA, F. Autopoiesis and Cognition. **Boston Studies in the Philosophy and History of Science**, vol. 42. Springer, Dordercht, 1980.

MATURANA, H.; VARELA, F. **A Árvore do Conhecimento – As Bases Biológicas da Compreensão Humana**. São Paulo: Palas Athena, 2019.

MAYER-SCHÖNBERGER, V. E CUKIER, K. **Learning With Big Data – The Future of Education** (Kindle Single). NY: Harper Business, 2014.

MENARY, R. **The Extended Mind**. Londres: MIT Press Books, 2012.

MERLEAU-PONTY, M. **Fenomenologia da Percepção**. Rio de Janeiro: Livraria Freitas Bastos, 1971.

MINSKY, M. (1974). A Framework for Representing Knowledge. **MIT-AI Laboratory Memo**, 306. Disponível em <https://courses.media.mit.edu/2004spring/mas966/Minsky%201974%20Framework%20for%20knowledge.pdf>. Acesso em 22 de fevereiro de 2022.

MINSKY, M. **The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind**. New York: Simon & Schuster, 2007.

MORAVEC, H. **Mind Children – The Future of Robot and Human Intelligence**. Cambridge: Harvard University Press, 1988

MITCHELL, M. **Why is AI Harder Than We Think**. Disponível em <https://arxiv.org/abs/2104.12871>. Acesso em 20 de outubro de 2022. DOI: doi.org/10.48550/arXiv.2104.12871

NEWEN, A.; DE BRUIN, L.; GALLAGHER, S. **The Oxford Handbook of 4E Cognition**. Oxford: University Press, 2018.

NILSSON, N. J. **The Quest for Artificial Intelligence – A History of Ideas and Achievements**. Nova York: Cambridge University Press, 2010.

NILSSON, N. J. **Introduction to Machine Learning – An Early Draft of a Proposed Textbook**. Department of Computer Science, Stanford University, 1996. Disponível em <https://ai.stanford.edu/~nilsson/MLBOOK.pdf>. Acesso em 8 de junho de 2021.

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

NOË, A. **Varieties of Presence**. Cambridge: Harvard University Press, 2012.

O'NEIL, C. **Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy**. Nova York: Crown, 2016.

PEA, R. The Prehistory of the Learning Sciences. In: EVANS, M.; PACKER, M. J.; SAWYER, R. K. **Reflections on the Learning Sciences**. Cambridge University Press, 2016.

PESSOA, M. Teresa (2002). **Aprender a pensar como professor: a Teoria da Flexibilidade Cognitiva** (cap. III). In: PESSOA, M. T., Aprender a Pensar como Professor - contributo da metodologia de casos na promoção da flexibilidade cognitiva. [Tese de Doutoramento, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra].

PRESSEY, S. Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-Instruction, **The Journal of Psychology**, 29:2, 417-447, 1950. DOI: 10.1080/00223980.1950.9916043

ROBOTICS TECHNOLOGY – EFFECTORS. **Electronics Teacher**. Disponível em <https://www.electronicsteacher.com/robotics/robotics-technology/effectors.php>. Acesso em 20 de dezembro de 2021.

ROHDE, M. **Enaction, Embodiment, Evolutionary Robotics**. Paris: Atlantis Press, 2010.

ROLLA, G. **A Mente Enativa**. Porto Alegre: Editora Fi, 2021

ROLLA, G. e FIGUEIREDO, N. Rolla, G., & Figueiredo, N. (2021). Bringing forth a world, literally. **Phenomenology and the Cognitive Sciences**. <https://doi.org/10.1007/s11097-021-09760-z>

ROSENBERGER, R.; VERBEEK, P. P. **Postphenomenological Investigations – Essays on Human-Technology Relations**. London: Lexington Books, 2015.

ROWLANDS, M. **The New Science of Mind – From Extended to Embodied Phenomenology**. Londres: MIT Press Books, 2013.

SANTOS, E. J. R.; BANNELL, R. I.; LEPORACE, C. P. Screen Time, Temporality and (Dis)embodiment. In: OLIVEIRA, L. (Org.). **Managing Screen Time in an Online Society**. 1ed. Aveiro: IGI Global, 2019, v. 1, p. 46-77.

SCHANK, R.; EDELSON, D. A role for AI in education: Using technology to reshape education. **Journal of Artificial Intelligence in Education**, 1. 3-20, 1989.

SELWYN, N. **Should Robots Replace Teachers? AI and the future of Education**. Medford: Polity Press, 2019.

SETH, A. **From unconscious inference to the beholder's share: Predictive perception and human experience**. *European Review*, 27(3), 378-410, 2019. doi:10.1017/S1062798719000061

SEVERINO, A. J. **Metodologia do Trabalho Científico**. 23ª ed. revisada e atualizada. São Paulo: Editora Cortez, 2007

SCHEUTZ, M. (Ed.). **Computationalism: new directions**. Cambridge, MA: The MIT Press, 2002.

SHEETS-JOHNSTONE, M. Thinking in Movement: Further Analyses and Validations. In: **Enaction – Toward a New Paradigm for Cognitive Science**. Cambridge: MIT Press, 2010

SMITH, B. C. **The Promise of Artificial Intelligence: Reckoning and Judgement**. Cambridge: MIT Press, 2019.

SKINNER, B. F. **Tecnologia do Ensino**. Editora da USP: Brasil, 1958.

STANFORD ENCYCLOPEDIA OF PHILOSOPHY. **The Frame Problem**. 2016. Disponível em: <https://plato.stanford.edu/entries/frame-problem/>

STAPLETON, M. Enacting Education. **Phenomenology and the Cognitive Sciences** (2021) 20:887–913.

STAPLETON, M. Enacting Environments – From Unwelts to Institutions (no prelo). A ser publicado em KARYN, L. (Ed). **Knowers and Knowledge in East-West Philosophy: Epistemology Extended**, 2022. Disponível em https://www.academia.edu/44980544/Enacting_Environments . Acesso em 8 de novembro de 2022.

THE FRAME PROBLEM. **Stanford Encyclopedia of Philosophy**, fevereiro de 2004 e fevereiro de 2016. Disponível em <https://plato.stanford.edu/entries/frame-problem/>. Acesso em 2 de novembro de 2022.

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

STOCKMAN, C. e NOTTINGHAM, E. Surveillance Capitalism in Schools: What’s the Problem? **Digital Culture & Education**, 14(1), 1–15

THOMPSON, E. **Mind in Life. Biology, Phenomenology, and the Sciences of Mind**. Cambridge: Harvard University Press, 2007.

TURING, A. Computing Machinery and Intelligence. **Mind**, Volume LIX, Issue 236, Outubro 1950, Pages 433–460, <https://doi.org/10.1093/mind/LIX.236.433>

UEXKÜLL, J. von. **A Foray into the Worlds of Animals and Humans**. With a Theory of Meaning. Tradução de J. D. O’Neill. Minneapolis: University of Minnesota Press, [1934] 2010.

USO DE INTERNET, TELEVISÃO E CELULAR NO BRASIL. **Educa IBGE**. Disponível em <https://educa.ibge.gov.br/jovens/materias-especiais/20787-uso-de-internet-televisao-e-celular-no-brasil.html>. Acesso em 18 de novembro de 2022.

VARELA, F.; THOMPSON, E.; ROSCH, E. **The Embodied Mind – Cognitive Science and Human Experience**. Revised Edition. Cambridge: MIT PRESS, 2016.

VÉLIZ, C. **Privacy is Power - Why and how you should take control of your data**. London, UK: Penguin (Bantan Press), 2020.

VÉLIZ, C. Moral Zombies: Why Algorithms Are Not Moral Agents. **AI and Society** 36 (2): 487–97, 2021. <https://doi.org/10.1007/s00146-021-01189-x>.

VÉLIZ, C. **If AI Is Predicting Your Future, Are You Still Free?** Disponível em <https://www.wired.com/story/algorithmic-prophecies-undermine-free-will/>. Acesso em 23 de junho de 2022.

VICARI, R. M. Influências das Tecnologias de Inteligência Artificial no Ensino. **Estudos Avançados** 35(101): 73-84, 2021.

WATTERS, A. **Teaching Machines**. Cambridge: MIT Press, 2021.

WATTERS, A. Gordon Pask's Adaptive Teaching Machines. **Hack Education**, 28 de março de 2015. Disponível em <https://hackeducation.com/2015/03/28/pask>. Acesso em 10 de setembro de 2021.

WILLIAMSON, B. **Big Data in Education. The digital future of learning, policy and practice**. London: SAGE Publications, 2017.

WILLIAMSON, B.; EYNON, R. Historical threads, missing links, and future directions in AI in education. **Learning, Media and Technology**, 45:3, 223- 235, 2020. DOI: [10.1080/17439884.2020.1798995](https://doi.org/10.1080/17439884.2020.1798995)

WILES, J.; DARTNALL, T. **Perspectives on Cognitive Science – Theories, Experiments and Foundations** – Vol. II. Connecticut: Ablex Publishing, 1999.

ZAHAVI, D. (2019). **Phenomenology: the basics**. Routledge.

ZAHAVI, D. A fenomenologia e o(s) problema(s) da intersubjetividade. In BANNELL, R. I.; MIZRAHI, M.; FERREIRA, G. (Orgs.) **(Des)educando a educação: Mentos, Materialidades e Metáforas**. Tradução de Camila De Paoli Leporace. Rio de Janeiro: PUC-Rio, 2021.

ZEBROWSKI, R. L.; MCGRAW, E. Autonomy and Openness in Human and Machine Systems: Participatory Sense-Making and Artificial Minds. **Journal of Artificial Intelligence and Consciousness** 08 (02): 303–23, 2021. <https://doi.org/10.1142/s2705078521500181>.

ZUBOFF, S. **The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power**. New York: PublicAffairs, 2019, 691p. ISBN 9781610395694

Apêndice

A construção da tese: metodologia & caminho(s) de pesquisa¹⁶⁷

Foi adotada na tese a metodologia proposta por Antônio Joaquim Severino, professor titular aposentado de Filosofia da Educação (FEUSP), tanto para a leitura como para a produção de textos filosóficos. A metodologia de Severino se divide em cinco passos, e esses passos foram adotados, então, não apenas para a leitura dos textos para esta pesquisa, mas para construir a argumentação da tese. Eis uma descrição bastante resumida da metodologia:

Antes de tudo: delimitar uma unidade de leitura – capítulo, artigo, ensaio etc e fazer uma leitura superficial, tendo um primeiro contato com o texto; nesta etapa, deve-se anotar dúvidas e pesquisar um pouco sobre o autor e sua produção.

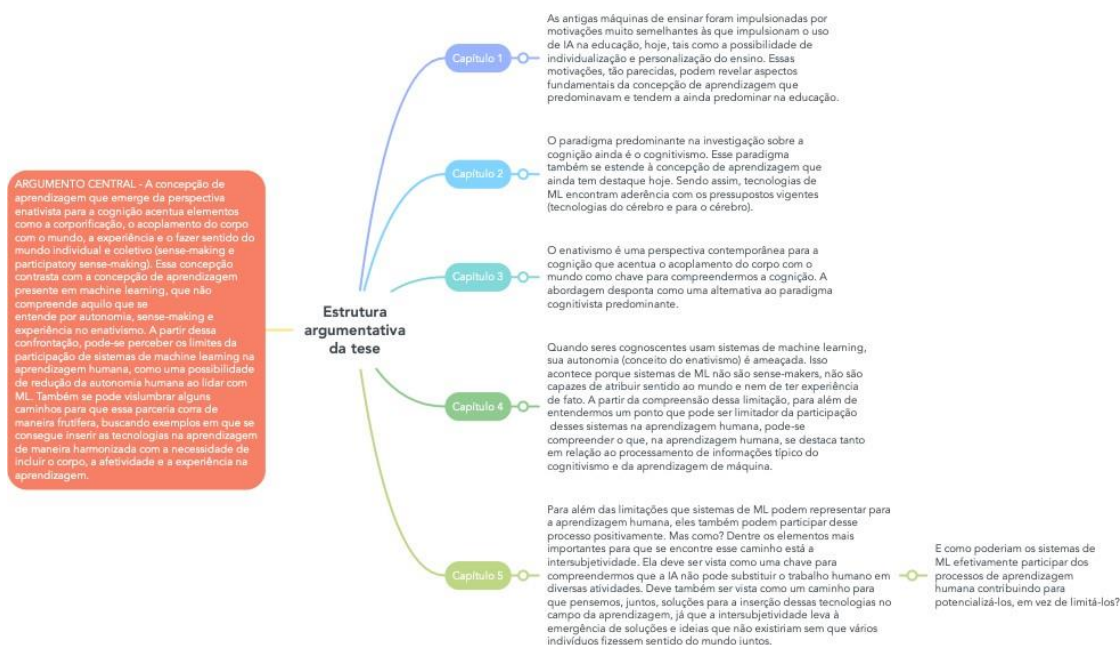
Os cinco passos:

- 1) Esquematização – Esta é a etapa em que se deve extrair as ideias principais do texto. Escrever na primeira pessoa é recomendado para facilitar a apropriação e compreensão dessas ideias. Não se deve emitir opinião nesta etapa, nem mesmo incluir informações oriundas de outros textos do mesmo autor;
- 2) Análise temática – Deve-se concluir qual é o tema, o problema tratado pelo autor ou pelos autores, qual o raciocínio feito para endereçar o problema e a tese a que se chega;

¹⁶⁷ Ao longo do meu percurso acadêmico no mestrado e no doutorado – em que, atenta a questões críticas envolvendo as tecnologias digitais na educação, trabalhei algumas dessas questões a partir da perspectiva da Filosofia da Educação – deparei-me muitas vezes com dúvidas sobre a própria construção de uma dissertação ou de uma tese teórica. Conversei muito com o professor Ralph Bannell, meu orientador na PUC-Rio, sobre esses questionamentos, e também com colegas mestrandos e doutorandos. No estágio à docência, alguns alunos da graduação quiseram saber mais sobre o processo e, me vendo instigada a explicá-lo, eu me punha a pensar sobre o percurso da pesquisa. Foi daí que pensei em acrescentar esta seção à tese: um pequeno anexo para compartilhar sobre a experiência de fazer uma tese em Filosofia da Educação.

- 3) Interpretação – O momento em que se inclui a visão de comentaristas e se pode situar aquele texto do autor no contexto mais amplo de sua produção e de suas ideias;
- 4) Problematização – O momento em que são levantados problemas para discussão, tanto textuais quanto de argumentação;
- 5) Síntese pessoal – Nesta etapa, reconstrói-se o texto a partir de reflexão pessoal; deve-se focar naquilo que, do argumento do autor, parece mais importante a quem está fazendo a análise do texto; deve-se por o autor em diálogo com outros.

Para desenvolver a argumentação da tese, cada capítulo foi estruturado a partir das etapas que compõem a análise temática. Isto é, cada um explicita um tema, o problema tratado, o raciocínio feito para endereçar o problema a tese alcançada em cada um. O mapa a seguir resume as etapas da metodologia (bastante simplificadas, uma vez que o mapa serve de base para o desenvolvimento):



A pesquisa teórico-filosófica não é sempre claramente compreendida, especialmente por conta da metodologia de pesquisa que envolve. Por vezes, um ensaio filosófico é confundido com uma revisão de bibliografia, a qual, apesar da

reconhecia importância, é diferente. A pesquisa de natureza teórico-filosófica tem uma função complementar às pesquisas de natureza empírica. A pesquisa teórico-filosófica se presta a investigar e trazer à tona perspectivas teóricas que poderão ser aplicadas futuramente em pesquisas empíricas. Com isso, faz emergir novas perguntas, que podem ajudar a encaminhar novas respostas ou novos tipos de respostas para perguntas de pesquisa já feitas.

A perspectiva enativista de que “o caminho se faz ao caminhar” adequa-se bastante bem à descrição do processo de desenvolver uma pesquisa, e mais ainda uma pesquisa em Filosofia; neste caso em particular, a Filosofia da Educação. Não se trata de *encontrar* um caminho já pronto, mas de encontrar uma *maneira de construir* essa trajetória e, por fim, dizer: enfim, caminhei, e este foi o caminho que construí. Vale ressaltar que a tese é uma etapa fundamental deste processo, uma espécie de balizadora para o trabalho, mas há inúmeras outras produções que acontecem no decorrer do caminhar e são igualmente importantes, como a participação em seminários e a produção de artigos e capítulos de livros. Tudo o que se faz nesse caminhar constrói o caminho. É preciso um trabalho constante de tecer, de costurar as ideias; e isso exige um olhar constantemente atento.

A pesquisa em geral, não apenas a filosófica, opera (ou deveria operar!) no terreno do inesperado, da surpresa. As escolhas feitas vão criando o caminho. Geram outras escolhas, abrem novas frentes, podem fechar algumas também (quando se percebe que uma determinada hipótese não faz sentido, por exemplo). Mas é fato que existem alguns elementos de referência para esse caminhar, que não devem ser desprezados, já que adotar uma metodologia de pesquisa é um compromisso tão importante quanto esse com o devir, com o *vir a ser* que subjaz à natureza da atividade do pesquisador. Para uma pesquisa em Filosofia da Educação, portanto de natureza teórica, é necessário alcançar um bom equilíbrio entre o comprometimento com a metodologia e a garantia de uma flexibilidade que permita que o pensamento emergja e o questionamento se mantenha constantemente. Talvez isso seja também comum e necessário a toda pesquisa; no entanto, no caso da pesquisa em Filosofia, esse é um dos **objetivos principais**: fazer emergir o pensar. É comum que a – ou as – pergunta(s) de pesquisa iniciais gerem outras perguntas, como se fossem subquestões relacionadas a essas questões principais, e isso é algo que exige que o pesquisador, de certa maneira, se acostume a lidar com a dúvida e a recortar o objeto de pesquisa constantemente.

Ter um projeto bem-organizado ajuda bastante nisso, pois se pode recorrer a ele sempre que necessário. Deve-se ter em mente que desenvolver uma pesquisa filosófica é mais sobre levantar questionamentos e colaborar para encaminhá-los do que propriamente sobre encontrar respostas. Mas, não se deve perder de vista as perguntas de pesquisa, motivadoras da tese, que o pesquisador precisa trabalhar para endereçar, ainda que admita que outras questões se abrirão (e que sua pesquisa dificilmente dará conta de encerrar o caso; pesquisar é contribuir para um campo, como pequenas gotas que compõem um oceano).

São atividades típicas de uma pesquisa com o perfil teórico-filosófico, entre outras: analisar premissas e pressupostos; investigar e delinear conceitos; trabalhar argumentos para construir outros – próprios do autor/pesquisador; contribuir para aproximar áreas de investigação; contribuir para transpor conceitos, teses, paradigmas de uma determinada área para outra de modo a preparar o caminho para que pesquisas empíricas sejam desenvolvidas. Parte-se de uma pergunta de pesquisa, levantam-se hipóteses para embasar os estudos; autores e abordagens são escolhidos inicialmente por algum motivo, são traçados os objetivos, a justificativa, como em todo projeto de pesquisa – mas, no caso de um projeto de pesquisa de fundamentação filosófica, permite-se que o desenvolvimento desse projeto seja elástico o suficiente para tomar novos rumos. Acredito que quem se propõe a fazer uma tese em Filosofia tem pelo menos duas perguntas para responder: a sua pergunta inicial de pesquisa (a questão da pesquisa) e mais uma: como adotar uma postura filosófica frente à sua questão de pesquisa?

Trata-se, então, de apropriar-se de determinado conhecimento, advindo de um certo campo ou de vários campos, com certos autores e abordagens, e de encaminhar uma resposta à pergunta com a qual o projeto de pesquisa se comprometeu. Esse é o conteúdo da pesquisa. Mas trata-se também de aprender a ter postura filosófica diante de uma questão.

A partir do momento em que um pesquisador em Filosofia da Educação se torna Doutor(a) – assim acredito – deverá levar essa postura consigo para toda a sua carreira ainda a ser trilhada. E ainda: lhe será requisitado que saiba despertar essa postura em seus orientandos, alunos, em seus parceiros de pesquisa. Esse é um dos papéis de um Doutor em Educação; assim eu acredito. Acredito, ainda, que essa postura envolve mais ler do que escrever, mais perguntar do que responder, e ainda: ouvir mais do que falar. Não se trata de não ter compromisso com as habilidades

didáticas que uma aula exige; nem de não ter consciência das inúmeras demandas de produção textual que a vida acadêmica traz; tampouco de se acomodar a uma dada posição de “gerador de questões” sem nunca tentar (ou ajudar a) encontrar possíveis respostas a elas. O que quero dizer é que a postura filosófica nos coloca em uma posição tal que nos conscientizamos de que, quanto mais se sabe, menos se sabe, como Sócrates diria; saímos da linha reta ensinar-aprender/professor-aluno; aprendemos a questionar as premissas; aprendemos a buscar outras variáveis – e a olhar embaixo de um tapete que às vezes ninguém quer levantar.

Em teses de Filosofia (e de Filosofia da Educação), não é tão comum que se descreva a metodologia utilizada, a qual aparece comumente apenas no projeto de pesquisa. Ainda assim, decidi incluir nesta seção esta breve explicação sobre como é a metodologia da pesquisa filosófica em educação por alguns motivos; o primeiro é que acredito ser válido mostrar como cheguei às conclusões e como estruturei eu pensamento no decorrer do processo de produção da tese. Além disso, acredito que esta breve descrição da metodologia poderá ajudar futuros mestrandos e doutorandos (em Filosofia da Educação ou em outras linhas), até quem sabe poderá colaborar para que acadêmicos já estabelecidos em suas linhas de pesquisa tenham contato com esse processo de tecer um trabalho filosófico e conheçam um pouco dos desafios enfrentados por aqueles que se lançam a esse tipo de pesquisa.