

A RESPOSTA QUE NÃO SE DEFENDE: THEUTH, THAMOUS E A INTELIGÊNCIA ARTIFICIAL GENERATIVA

THE ANSWER THAT CANNOT DEFEND ITSELF: THEUTH, THAMOUS AND GENERATIVE ARTIFICIAL INTELLIGENCE

LA RESPUESTA QUE NO SE DEFIENDE: THEUTH, THAMOUS Y LA INTELIGENCIA ARTIFICIAL GENERATIVA

Narda Teles Yamane¹
Rozineide Iraci Pereira da Silva²

RESUMO: Este artigo examina em que medida a crítica à escrita apresentada no *Fedro*, de Platão, permite interrogar a inteligência artificial generativa e suas implicações para a aprendizagem. Desenvolve-se uma pesquisa bibliográfica, de abordagem hermenêutico-crítica, centrada nas passagens 274c–276e da edição bilíngue traduzida por José Cavalcante de Souza. A análise distingue o juízo de Thamous sobre memória e aparência de sabedoria das considerações posteriores de Sócrates acerca do discurso escrito. Nestas, identifica três limitações: a invariância diante da interrogação, a circulação sem discernimento do destinatário e a impossibilidade de defesa autônoma. O confronto com a literatura sobre sistemas generativos mostra que a produção de réplicas variáveis modifica a primeira limitação, enquanto a adequação pedagógica e a justificação permanecem exigências que não se resolvem pela fluência da resposta. Sustenta-se que corrigir, discordar ou reformular um enunciado não demonstra, isoladamente, conhecimento nem dispensa a responsabilidade humana. Conclui-se que o problema educacional se desloca da capacidade de a máquina responder para as condições em que o aluno compreende, examina e sustenta aquilo que apresenta como saber.

1

Palavras-chave: *Fedro*. inteligência artificial generativa. Aprendizagem.

ABSTRACT: This article examines the extent to which the critique of writing presented in Plato's *Phaedrus* can inform an inquiry into generative artificial intelligence and its implications for learning. It develops a bibliographical study with a hermeneutic-critical approach, focusing on passages 274c–276e in the bilingual edition translated by José Cavalcante de Souza. The analysis distinguishes Thamous's judgment on memory and the appearance of wisdom from Socrates's subsequent discussion of written discourse. It identifies three limitations in the latter: invariance when questioned, circulation without discerning the addressee, and inability to defend itself autonomously. Comparison with research on generative systems indicates that variable replies modify the first limitation, whereas pedagogical appropriateness and justification remain requirements that fluent responses cannot resolve on their own. The article argues that correcting, disagreeing, or reformulating a statement does not, by itself, demonstrate knowledge or remove human responsibility. It concludes that the educational question shifts from the machine's capacity to answer to the conditions under which students understand, examine, and support what they present as knowledge.

Keywords: *Phaedrus*. Generative artificial intelligence. Learning.

¹Discente do Curso de Doutorado em Ciências da Educação da Christian Business School-CBS.

²Professora do Ensino Superior e professora Orientadora da Christian Business School-CBS Ph.D. Doutora em Ciências da Educação, Mestra em Ciências da Educação pela Universidade Federal de Alagoas-UFAL, Psicopedagoga, Pedagoga, Analista do Comportamento Aplicada, Especialista em Escrita Acadêmica Avançada.

RESUMEN: Este artículo examina en qué medida la crítica de la escritura presentada en el *Fedro*, de Platón, permite interrogar la inteligencia artificial generativa y sus implicaciones para el aprendizaje. Se desarrolla una investigación bibliográfica, de enfoque hermenéutico-crítico, centrada en los pasajes 274c–276e de la edición bilingüe traducida por José Cavalcante de Souza. El análisis distingue el juicio de Thamous sobre la memoria y la apariencia de sabiduría de las consideraciones posteriores de Sócrates acerca del discurso escrito. En estas se identifican tres limitaciones: la invariancia ante la interrogación, la circulación sin discernimiento del destinatario y la imposibilidad de defensa autónoma. El contraste con la literatura sobre sistemas generativos muestra que la producción de réplicas variables modifica la primera limitación, mientras que la adecuación pedagógica y la justificación siguen siendo exigencias que no se resuelven mediante la fluidez de la respuesta. Se sostiene que corregir, discrepar o reformular un enunciado no demuestra por sí solo conocimiento ni elimina la responsabilidad humana. Se concluye que el problema educativo se desplaza hacia las condiciones en que el estudiante comprende, examina y sostiene aquello que presenta como saber.

Palabras clave: *Fedro*. inteligencia artificial generativa. Aprendizaje.

I INTRODUÇÃO

A importância de uma coisa não se mede com fita métrica nem com balanças nem barômetros etc. A importância de uma coisa há que ser medida pelo encantamento que a coisa produza em nós.
— Manoel de Barros

A citação de Manoel de Barros diz algo que a academia costuma esquecer: a importância de uma coisa não se esgota em sua função, mas no que ela opera em nós. Transportada para o campo da aprendizagem, a intuição do poeta ganha uma aresta incômoda. Um aluno pode dispor de um texto que responde à atividade proposta — responde bem, com clareza, com encanto — e, no entanto, não conseguir explicar por que uma afirmação está ali, de que depende sua validade ou como ela precisaria mudar diante de uma objeção pertinente. O encantamento da resposta não é a compreensão da resposta.

Trata-se de uma possibilidade analítica, não de um relato de campo produzido por esta pesquisa. A mesma dificuldade pode ocorrer com textos copiados de livros, materiais preparados por terceiros ou discursos decorados. A novidade da inteligência artificial generativa não está em ter inventado a distância entre exposição e compreensão, mas em fornecer respostas adicionais dentro da própria situação em que essa distância seria interrogada.

O *Fedro* oferece, a nosso ver, um ponto de partida rigoroso para examinar o problema. No relato de Theuth e Thamous, a escrita é submetida a uma avaliação que distingue as intenções do inventor dos efeitos atribuídos à invenção. Depois do relato, Sócrates pergunta pelo comportamento do discurso escrito quando alguém deseja aprender e o interroga, identificando três limitações cujo exame detalhado ocupa a fundamentação teórica deste artigo.

O movimento do diálogo permite perguntar não apenas o que uma técnica faz circular, mas o que acontece com aquele que dela se serve.

A aproximação entre o *Fedro* e as tecnologias contemporâneas não é inédita. Aguiar (2023, p. 184, 186) já inclui o ChatGPT em sua reflexão e discute a diferença entre comunicação recíproca e circulação de informações. Não reivindicamos, portanto, descobrir uma relação ausente da literatura. É nossa intenção propor um recorte mais delimitado: distinguir as limitações atribuídas por Sócrates à escrita — a invariância diante da interrogação, a circulação sem discernimento do destinatário e a impossibilidade de defesa autônoma — e examinar o que se modifica quando o dispositivo passa a produzir réplicas variáveis. A pergunta que orienta o artigo é esta: em que sentido a capacidade de responder altera a crítica platônica à escrita, e quais exigências de conhecimento e formação não se resolvem apenas por essa alteração?

Pensamos que a responsividade técnica modifica a comparação com o escrito fixo, mas não resolve, por si, a adequação do discurso ao aluno, a validade de suas razões nem a apropriação do conhecimento. O título "A resposta que não se defende" designa esse problema, e não uma incapacidade literal de a inteligência artificial generativa produzir contra-argumentos. Um sistema pode contestar, corrigir e apresentar justificativas. A questão é saber o que autoriza essas justificativas e quem as examina e assume quando elas passam a orientar o estudo, a avaliação ou uma decisão. O retorno a Platão interessa precisamente onde a analogia deixa de ser confortável e precisa ser reconstruída.

2 O FEDRO E A QUESTÃO DA TÉCNICA

2.1 O *Fedro* como Texto sobre a Técnica

Ler o *Fedro* como texto sobre a técnica exige preservar a situação em que a escrita é discutida. O diálogo não começa por um tratado sobre instrumentos. Sua investigação atravessa discursos sobre o amor, a alma, a retórica e as condições do dizer. Nas passagens próximas ao mito, a questão é a conveniência e a inconveniência de escrever, depois de examinadas as exigências de uma arte dos discursos. O objeto em julgamento é uma prática que interfere na relação entre palavra, conhecimento e interlocução, e não um artefato isolado de seus usos (Platão, 2016, p. 187–191, 273d–274c).

Sócrates introduz a narrativa como algo ouvido dos antigos. A distância narrativa importa: Platão escreve um diálogo no qual Sócrates conta uma história e, dentro dela, duas personagens avaliam uma invenção. Apagar esses níveis faria parecer que todas as proposições

pertencem a uma única voz. Apagar esses níveis seria reduzir Platão a um manual. A nosso ver, a análise precisa distinguir a defesa de Theuth, a advertência de Thamous e a elaboração socrática que se segue. São movimentos articulados, mas não intercambiáveis.

Seria precipitado concluir que um diálogo escrito simplesmente condena toda escrita. O próprio texto exige do leitor o trabalho de acompanhar objeções, reconhecer mudanças de posição e relacionar passagens. Vale dizer que o diálogo conserva um lugar para os registros escritos como lembrança e exercício, sem lhes atribuir suficiência para ensinar a verdade. Quando Sócrates distingue o discurso que se planta "com ciência na alma" daquele que apenas se fixa em papiro ou pedra, não proíbe a escrita — interroga as condições sob as quais ela participa do saber e as condições sob as quais se apresenta como sua substituta. Nas palavras do próprio Sócrates:

Mas creio que muito mais belo é o empenho em torno disso quando alguém, usando a arte dialética, pega uma alma condigna e nela planta e semeia com ciência discursos que a si mesmos e ao que plantou [277a] são capazes de assistir, e, não estéreis, têm uma semente da qual outros, em outros naturais brotando, são capazes de sempre garantir imperecível este efeito, e fazendo no que os tem a maior felicidade que é possível ao homem (Platão, 2016, p. 201, 277a).

A tensão não se resolve afirmando que Platão seria incoerente por escrever. O que se segue obriga a pensar quando a escrita serve ao saber e quando toma o seu lugar.

4

Rodrigues (2015) mostra que a crítica do *Fedro* não equivale à recusa absoluta da escrita. Sua leitura destaca a aprendizagem e a adequação do discurso a quem aprende, inclusive quanto ao momento e à extensão da exposição. Esse antecedente permite situar nosso problema: a insuficiência de um suporte não elimina seus usos, mas impede que se confunda a conservação de enunciados com a formação de quem poderá compreendê-los.

Ao trazer esse problema para a IAG, não afirmamos que os dispositivos sejam idênticos. O manuscrito antigo e o assistente conversacional diferem material e funcionalmente. Uma aproximação produtiva deve explicitar as diferenças, sobretudo a possibilidade de gerar respostas ao longo da interação. A pergunta filosófica não depende de ocultá-las. Depende de examinar se a transformação do comportamento visível do discurso altera igualmente as condições pelas quais ele pode valer como conhecimento.

2.2 O *Phármakon* e a Recepção do *Fedro* na Filosofia da Técnica

A palavra *phármakon* — em grego, ao mesmo tempo remédio e veneno — ocupa um lugar decisivo na disputa sobre a escrita. Na tradução aqui utilizada, Theuth anuncia um

medicamento para a memória e a sabedoria, enquanto Thamous contesta o benefício anunciado. A invenção recebe valorizações diferentes conforme se considera aquilo que promete conservar e aquilo cujo exercício pode dispensar. O desacordo já impede tratar a técnica como simples acréscimo exterior, incapaz de modificar a atividade de quem a utiliza. Thamous formula o desacordo em dois momentos — separa quem engendra a arte de quem lhe avalia os efeitos e, logo a seguir, distingue memória de recordação e a sabedoria de sua aparência (Platão, 2016, p. 193, 274e–275a) —, passagens que examinamos de perto adiante (seções 2.4 e 2.5).

Na recepção discutida por Aguiar (2023), a leitura derridiana do *phármakon* e seu desdobramento farmacológico em Stiegler permitem pensar a não neutralidade da técnica e a ambivalência de seus efeitos. Essa abordagem também considera a reciprocidade comunicativa; não seria correto reduzi-la a uma enumeração de vantagens e desvantagens nem afirmar que ela ignora a interlocução. Nosso exame acrescenta uma discriminação analítica entre invariância, adequação e defesa, sem pretender refutar essa tradição.

Há, nesse ponto, uma precaução metodológica. A ambivalência não pode funcionar como conclusão pronta: declarar que a tecnologia pode ajudar ou prejudicar ainda não esclarece o problema. É necessário especificar qual atividade é apoiada, qual operação passa a ser delegada e em que condições o apoio favorece ou dispensa a aprendizagem. Essas perguntas orientam a interpretação desenvolvida adiante. Elas não são apresentadas como resultados experimentais nem como propriedades invariáveis de qualquer uso da IAG.

5

O deslocamento proposto consiste em examinar separadamente o que acontece quando o discurso é solicitado a responder, quando precisa considerar um destinatário e quando suas razões são contestadas. A distinção impede que o sucesso em uma dessas operações seja tomado como demonstração das demais. Uma réplica pertinente pode coexistir com uma razão falsa; uma razão correta pode ser pedagogicamente inadequada; uma exposição adequada pode ser recebida sem elaboração. A análise do *Fedro* oferece a ocasião para tornar essas diferenças visíveis.

Por isso, o problema não será resolvido escolhendo previamente o remédio ou o veneno. O que nos importa é identificar as mediações pelas quais uma capacidade técnica passa a ser tratada como capacidade formativa. A própria rapidez pode ter sentidos diferentes: liberar tempo para investigar ou eliminar a investigação; esclarecer uma dúvida ou encerrar prematuramente a pergunta. Nenhum desses efeitos deve ser presumido apenas pela presença do dispositivo.

2.3 Linguagem, Experiência e Formação de Si

No campo educacional, a existência de um discurso correto não dispensa a pergunta por sua relação com quem aprende. Quando utilizamos aqui a expressão formação de si, não designamos um desenvolvimento isolado dos outros, nem uma exigência de produzir pensamentos sem qualquer mediação. Referimo-nos à transformação da relação que alguém estabelece com o que sabe, com o que ainda não compreende e com as razões que consegue oferecer. O interesse do artigo permanece delimitado à aprendizagem e à justificação do discurso.

Larrosa (2002) distingue a disponibilidade de informação da experiência que afeta e transforma alguém. Sua reflexão não constitui uma teoria experimental dos efeitos da IAG, mas ajuda a formular uma cautela: receber conteúdos e ter algo a dizer sobre eles não assegura que se tenha alterado a própria relação com o que foi recebido. A questão formativa requer atenção ao modo como o sujeito se envolve, e não somente ao volume de informações ao qual tem acesso.

Back (2023), em diálogo com a hermenêutica de Gadamer, diferencia compreensão e aprendizagem, destacando nesta a capacidade de realizar a atividade relacionada ao que foi compreendido. Essa distinção é fecunda para nosso recorte. Acompanhar uma explicação, reconhecer sua plausibilidade e conseguir reconstruir seu caminho são acontecimentos que não convém identificar antecipadamente. Não se trata de impor uma sequência psicológica universal, mas de impedir que um único indício de desempenho represente todo o processo de aprender.

Podemos então delimitar três planos. No primeiro, uma resposta é produzida e circula. No segundo, suas afirmações são examinadas quanto à validade e à pertinência. No terceiro, alguém passa a operar com o que compreendeu, reformulando, aplicando e revendo razões. Os planos podem se entrelaçar, mas a passagem de um a outro precisa ser investigada. A produção de um texto não atesta sozinha sua validade; sua validade não atesta sozinha a aprendizagem do aluno que o apresenta.

O tempo da formação aparece nesse intervalo. Não como uma duração previamente calculável, nem como elogio da dificuldade pela dificuldade, mas como possibilidade de voltar ao que não se compreendeu, experimentar uma formulação e reconhecer por que ela falha. Há ajudas que tornam essa elaboração possível. Há atalhos que apenas a dispensam. A distinção

que procuramos elaborar não opõe aprender sozinho a aprender com apoio; pergunta se o apoio preserva para o aluno alguma atividade intelectual que ele precisará efetivamente realizar.

2.4 O Juízo sobre a Invenção

Theuth apresenta a escrita a partir do benefício que espera produzir. Sua defesa associa o domínio das letras ao aumento da memória e da sabedoria dos egípcios. Thamous interrompe essa associação. Não nega que uma invenção tenha sido realizada; contesta que seu inventor detenha, por esse fato, o juízo suficiente sobre aquilo que ela fará aos seus usuários. A resposta do rei explicita a diferença:

“Ó tecnicíssimo Theuth, um é o capaz de engendrar os elementos da arte, outro o de julgar a parte de dano e de utilidade que ela tem para os que vão usá-la.” (Platão, 2016, p. 193, 274e).

O juízo introduz uma mudança de perspectiva. A competência de produzir a arte não se identifica com a capacidade de avaliar sua participação na vida de quem a aprenderá. Essa diferença não autoriza presumir que o inventor esteja necessariamente enganado ou que quem observa de fora seja, por princípio, melhor juiz. No relato, ela impede que a afeição pela obra se converta em garantia de seu valor. Uma promessa precisa responder pelos efeitos que anuncia.

7

Thamous dirige sua avaliação aos alunos. A pergunta não é apenas se as letras conservam alguma coisa, mas o que acontece à memória daqueles que passam a confiar nelas. O benefício alegado por Theuth é reexaminado a partir da atividade humana que deveria favorecer. Esse é o primeiro resultado da leitura: a técnica é julgada em sua relação com uma prática, e a existência de um produto não decide antecipadamente a qualidade dessa relação.

O texto também permite distinguir intenção e justificação. Theuth pode desejar beneficiar os egípcios e, ainda assim, descrever inadequadamente os efeitos da invenção. A boa intenção não elimina a necessidade do exame; tampouco o exame precisa atribuir uma intenção maléfica ao inventor. O desacordo ganha interesse precisamente porque não se reduz à oposição entre personagens bem-intencionadas e mal-intencionadas. Está em questão a passagem, não demonstrada, da potência da técnica à melhoria da aprendizagem daquele que a utiliza.

2.5 A Troca de Memória por Recordação

O ponto seguinte do julgamento explicita o que Thamous considera uma troca indevida. O inventor chama de fortalecimento da memória aquilo que o rei descreve como dependência de sinais externos. Na tradução utilizada, memória e recordação não são palavras substituíveis:

Pois isto, nos que o aprenderam, esquecimento em suas almas produzirá com o não exercício da memória, porque, na escrita confiando, é de fora, por alheias impressões e não por eles mesmos, que se recordam; assim, não para memória, mas para recordação achaste um medicamento. E, da sabedoria, aos teus aprendizes transmites uma aparência, não a verdade (Platão, 2016, p. 193, 275a).

O que está sob suspeita não é a presença de um registro exterior considerada isoladamente. É a confiança nele quando passa a dispensar o exercício que deveria apoiar. O discurso de Thamous atribui à escrita a capacidade de relembrar por intermédio de marcas, mas recusa que essa operação seja automaticamente chamada de memória ou de sabedoria. A distinção entre *mnēmē* (a memória que se exerce por si) e *hypómnēsis* (a recordação por meio de marcas externas), presente no texto grego, orienta essa passagem; não a converteremos em uma teoria geral da memória ou da reminiscência platônica.

A aparência de sabedoria decorre de outra identificação: estar informado poderia fazer alguém parecer instruído. A advertência não autoriza chamar de ignorante qualquer pessoa que utilize um recurso de registro. Incide sobre a pretensão de apresentar como saber aquilo que não passou pela aprendizagem correspondente. O perigo descrito no mito é uma relação alterada com o próprio não saber: a pessoa já dispõe de enunciados suficientes para que sua necessidade de aprender deixe de aparecer.

Convém manter a natureza do argumento. Thamous formula um juízo no interior de uma narrativa filosófica, não um resultado obtido em experimento sobre memória. Sua advertência oferece uma questão para a pesquisa educacional; não dispensa essa pesquisa. A passagem não demonstra que o recurso externo necessariamente empobreça todas as aprendizagens. Permite, antes, perguntar quando o apoio à recordação se apresenta como substituto de uma capacidade que ainda precisaria ser exercitada.

Ao final do relato, a verdade continua a exigir exame, independentemente do prestígio ou da origem de quem fala. Quando Fedro chama a atenção para a facilidade de Sócrates em criar histórias, a resposta socrática recorda a disposição de ouvir até um carvalho ou uma pedra, desde que dissessem a verdade (Platão, 2016, p. 193, 275b–275c). Essa passagem dificulta uma

leitura que rejeitasse uma proposição apenas por sua procedência. O problema não se resolve perguntando somente quem a produziu. É preciso examinar se aquilo que se diz se sustenta.

2.6 As Três Marcas do Discurso Escrito

Convém retomar o critério com que Sócrates examina a escrita depois do relato de Thamous, porque é dele que depende o que se segue. A objeção não incide apenas sobre o conteúdo do que se escreve, e sim sobre o modo como o escrito se comporta depois de posto no mundo. A comparação com a pintura introduz a aparência de vida e, em seguida, a dificuldade que surge quando alguém se dirige ao discurso querendo aprender. A passagem em que Sócrates descreve o que de terrível tem a escrita merece ser transcrita na íntegra, porque constitui o objeto central da nossa análise:

O que de terrível sem dúvida, ó Fedro, tem a escrita é realmente a sua semelhança com a pintura. E de fato os seres que esta engendra estão como se fossem vivos; porém se lhes perguntas algo, solene e total é o seu silêncio. E o mesmo fazem também os discursos; poderias crer que um pensamento anima o que eles dizem; mas se algo perguntas do que é dito, querendo aprender, uma só coisa apenas eles indicam e a mesma sempre. E uma vez [275e] escrito, fica rolando por toda parte todo discurso, igualmente entre os que sabem como entre aqueles com os quais nada tem a ver, e nunca sabe a quem ele justamente deve falar e a quem não. Transgredido e não com justiça censurado, do pai sempre ele precisa como assistente; pois ele próprio não é capaz nem de se defender nem de se assistir por si mesmo (Platão, 2016, p. 195, 275d–275e).

9

São três marcas, e não uma só: a invariância da resposta (“uma só coisa apenas eles indicam e a mesma sempre”), a incapacidade de discernir o destinatário (“nunca sabe a quem ele justamente deve falar e a quem não”) e a impossibilidade da defesa (“não é capaz nem de se defender nem de se assistir por si mesmo”). Essa tríade é uma organização analítica que propomos para a passagem, não uma lista numerada por Platão. A primeira concerne à resposta diante da interrogação; a segunda, à circulação e à pertinência do dizer; a terceira, à assistência de que o discurso necessita quando é contestado. Distinguem-se para que não se tome uma delas como equivalente suficiente das outras.

A expressão “querendo aprender” situa a objeção em uma relação precisa. Sócrates não pede ao escrito que varie por capricho. A repetição torna-se um limite porque há uma pergunta que deseja compreender algo do que foi dito. Também a circulação não é examinada apenas como alcance material. O escrito chega a destinatários diferentes sem, por si, avaliar o que dizer a cada um. Por fim, a necessidade do pai aparece diante da censura injusta. A defesa é uma resposta à contestação, e não a conservação obstinada de qualquer enunciado.

Poucas linhas adiante, Sócrates opõe ao discurso escrito outro, que chama de irmão legítimo deste: o que se escreve com ciência na alma de quem aprende, capaz de defender-se e de saber diante de quem falar e diante de quem calar. E Fedro conclui que o escrito, comparado a esse, é apenas o seu simulacro. A definição socrática é breve: "O que se escreve com ciência na alma do que aprende, e que pode se defender e sabe falar e calar diante de quem é preciso" (Platão, 2016, p. 197, 276a).

Segue-se um segundo resultado interpretativo. O contraste não pode ser reduzido à presença de uma resposta verbal. Em 276a, saber falar, calar e defender o discurso está vinculado à inscrição com conhecimento, retomada por Fedro como discurso daquele que sabe. A palavra ciência traduz aqui *epistēmē*, sem equivalência imediata com o sentido moderno de ciência experimental. Do mesmo modo, alma deve ser preservada no horizonte do diálogo, sem substituição automática por consciência computacional, mente individual ou identidade contemporânea.

Assim, a exterioridade das três marcas não as torna irrelevantes, mas impede usá-las como testes autossuficientes. Produzir variações, reconhecer destinatários ou resistir a uma censura não prova, isoladamente, que se satisfaça a exigência de saber. A leitura não oferece um protocolo de detecção da inteligência. Mostra que o comportamento do discurso precisa ser relacionado à compreensão que o sustenta. É essa relação que o confronto com os sistemas generativos deverá interrogar, sem supor que a mera mudança do suporte a tenha resolvido.

10

3 MÉTODOS

3.1 Natureza e Delimitação da Pesquisa

Este artigo examina se a crítica à escrita desenvolvida no relato de Theuth e Thamous e na discussão socrática que o sucede ainda alcança a inteligência artificial generativa. Trata-se de pesquisa qualitativa, de natureza bibliográfica e abordagem hermenêutico-crítica, no campo da Filosofia da Educação.

O mito não entra como ilustração de um problema formulado por outros meios. Ele interessa porque submete uma técnica de linguagem a juízo. A análise acompanha o relato e seu desdobramento no diálogo, concentrando-se no modo como o discurso se comporta quando alguém o interroga. É esse problema, e não uma explicação geral da obra platônica, que constitui o objeto da investigação.

Daí a delimitação das passagens. Em 274c–275b, a invenção é apresentada e avaliada, e o julgamento separa quem engendra a arte de quem considera os efeitos de seu uso. Em 275d–275e, Sócrates examina a repetição do escrito, sua circulação e sua dependência de assistência. Em 276a–276b, contrapõe-lhe o discurso inscrito com ciência na alma de quem aprende. A leitura estende-se até 276e para não separar essa oposição da analogia agrícola e dos usos da escrita que o próprio texto admite.

3.2 Fontes Bibliográficas

Como a análise recai sobre termos determinados, a escolha da edição importa. Adotamos a edição bilíngue da Editora 34, em tradução de José Cavalcante de Souza, publicada em 2016. As citações foram conferidas no exemplar digital utilizado, preservando a redação portuguesa e distinguindo sua paginação impressa da contagem de páginas do arquivo. As chamadas reúnem autor, ano, página impressa e localização Stephanus. Não se reivindica cotejo sistemático de traduções nem se atribuem à tradução portuguesa conclusões filológicas que exigiriam investigação própria.

Por essa mesma razão, as passagens decisivas são reproduzidas em citação direta: constituem o objeto analisado, e não apenas apoio de autoridade ao argumento. A bibliografia complementar foi selecionada intencionalmente pela pertinência a três problemas: a interpretação da escrita no *Fedro*; a relação entre compreensão e aprendizagem; e as possibilidades e limitações documentadas dos sistemas generativos. Foram consultados textos de periódicos, anais, repositórios acadêmicos e o guia da UNESCO. Não se trata de revisão sistemática ou levantamento exaustivo da produção científica.

3.3 Procedimentos e Limites da Análise

O procedimento compreende a identificação de quem enuncia cada objeção, a distinção de seus critérios e o confronto conceitual com a literatura contemporânea. Chamamos responsividade técnica à capacidade de produzir uma réplica condicionada pela solicitação e pelo contexto disponível. Essa definição não equivale a uma definição de inteligência, conhecimento ou consciência. Adequação pedagógica e justificação são examinadas separadamente, evitando concluir uma capacidade a partir da simples presença de outra.

Os resultados apresentados na seção seguinte são interpretativos. Não foram coletados depoimentos, trabalhos escolares ou interações próprias com modelos para constituir um corpus

empírico. Estudos experimentais citados na discussão conservam seus contextos e limites; não funcionam como comprovação retrospectiva de Platão. As propostas educacionais que encerram o percurso são decorrências argumentativas a serem investigadas, e não intervenções cuja eficácia este artigo teria demonstrado.

3.4 Uso de ferramentas de IA

Durante a preparação deste manuscrito, a autora utilizou o Chat GPT 5.4 para revisão gramatical e ortográfica. A ferramenta não foi utilizada para geração de conteúdo, análise de dados ou interpretação de resultados. A autora revisou e assume responsabilidade integral pelo trabalho.

4 RESULTADOS E DISCUSSÃO

4.1 A Marca que Caiu

Diante da inteligência artificial generativa, a invariância deixa de caracterizar a interação do mesmo modo que caracteriza a página fixa. Um assistente conversacional pode ajustar o vocabulário à solicitação, mudar o registro, retomar elementos disponíveis no contexto e produzir uma nova formulação. Isso não implica que toda réplica seja diferente ou pertinente, mas impede descrevê-lo simplesmente como um escrito que repete sua formulação original. Estudos sobre esses sistemas documentam tanto sua adaptação às intervenções quanto mudanças inadequadas decorrentes delas (Sharma *et al.*, 2024).

Se a crítica socrática se resumisse à invariância, essa diferença pareceria suficiente para afastá-la. Seria uma conclusão precipitada. O que caiu foi a possibilidade de identificar ausência de interlocução apenas pela repetição material do mesmo enunciado. Não caiu a necessidade de examinar o que a réplica acrescenta, por quais razões ela muda e em que medida permite que alguém aprenda. O sucesso conversacional abre a pergunta, não a encerra.

Há também uma diferença entre o enunciado gerado e o sistema que o gera. Uma mensagem já produzida permanece um registro, passível de leitura e circulação; a resposta seguinte é produzida por um sistema em funcionamento. A variabilidade não pertence a uma página que teria adquirido vida própria, mas a uma organização técnica capaz de produzir novos enunciados. Quando se atribui à mensagem toda a capacidade do sistema, ou ao sistema uma unidade de compromisso semelhante à de uma pessoa, etapas diferentes da interpretação são fundidas.

Essa diferença importa para a educação. Uma explicação alternativa pode oferecer a passagem que faltava ao aluno; pode também substituir uma dificuldade ainda não examinada por outra formulação igualmente pouco compreendida. Não há contradição em reconhecer o valor da explicação e negar que ela assegure aprendizagem. O que precisa ser investigado é a relação entre a intervenção e a atividade do aluno. A simples multiplicação de respostas não informa se houve essa atividade.

Por isso, a oposição decisiva não será entre imobilidade e movimento do texto. Uma conversa humana pode ser estéril apesar de suas réplicas, enquanto um livro pode desencadear uma investigação duradoura. O critério não pode ser o número de turnos nem a aparência de atenção. A pergunta é se a interação permite elaborar razões que não se reduzem à reprodução daquilo que acabou de ser recebido. O diagnóstico é relacional: considera o dispositivo, a tarefa, o aluno e a mediação que os aproxima.

5.2 As Duas que Permanecem

As outras duas marcas permanecem como exigências a examinar, não como incapacidades técnicas que se conservariam intactas. Seria inexato afirmar que a IAG responde indistintamente a qualquer pedido. Bai *et al.* (2022), em pesquisa apresentada como preprint, descrevem um procedimento de treinamento orientado por princípios que permite recusar solicitações e explicar recusas. Esse resultado não equivale a atribuir juízo moral humano ao sistema, mas basta para tornar inadequada a afirmação de que ele não pode recusar ou contestar.

A questão da adequação exige, portanto, uma formulação mais cuidadosa. Adaptar um texto à idade declarada, ao nível solicitado ou à forma da pergunta não assegura que a intervenção corresponda ao que aquele aluno precisa aprender. A adequação pedagógica envolve finalidades que nem sempre coincidem com a satisfação imediata do pedido. Em certas tarefas, a melhor ajuda pode ser uma pergunta, uma pista ou o adiamento da resposta completa. Essa escolha depende de critérios educacionais que precisam ser definidos e avaliados, ainda que parte de sua execução seja automatizada.

Uma pessoa tampouco conhece perfeitamente o outro. O professor interpreta sinais, pode errar e precisa rever seu julgamento. Não contrapomos, assim, uma máquina incapaz a um educador infalível. O que defendemos é que as condições e os limites da intervenção permaneçam discutíveis por quem responde pela prática educativa. Saber a quem falar, nessa

aproximação com o *Fedro*, deixa de ser uma propriedade abstrata do dispositivo e passa a designar uma responsabilidade situada na organização do ensino.

A terceira marca exige outra distinção. Defender-se não é reformular-se. Mas defender-se também não é conservar uma posição contra toda objeção. Quem reconhece um erro pode justificar por que o abandona; quem insiste pode apenas repetir uma crença. A racionalidade de uma revisão depende das razões que a motivam, e não do desconforto que ela produz ou da resistência pessoal em aceitá-la. Seria igualmente equivocado tomar a concessão como prova de ausência de pensamento e a obstinação como prova de conhecimento.

Sharma *et al.* (2024) encontraram, em cinco assistentes e nas tarefas examinadas, comportamentos de concordância indevida com o usuário, incluindo alterações de respostas corretas após contestação. O estudo oferece uma advertência delimitada: a adaptação conversacional pode favorecer a concordância em detrimento da correção. Não demonstra que todos os sistemas sempre cedem nem que qualquer revisão automatizada seja inválida. Para nosso argumento, mostra por que a presença de uma réplica não basta para avaliar a qualidade da defesa.

O contrário também precisa ser admitido. Se um sistema apresenta uma demonstração válida ou uma objeção pertinente, sua origem não torna o argumento falso. O exame deve incidir sobre premissas, evidências e inferências. Contudo, a possibilidade de gerar uma boa razão não dispensa a verificação de sua pertinência naquele caso, assim como a produção de uma justificativa verbal não prova que ela descreva fielmente o processo que produziu a resposta. A avaliação do argumento e a atribuição de conhecimento ao sistema são problemas relacionados, mas distintos.

Chegamos, então, ao sentido mais preciso da defesa neste artigo. Uma resposta não se torna garantia de si porque consegue gerar outra resposta em seu auxílio. Ela precisa permanecer aberta ao confronto com fontes, argumentos e situações que não são estabelecidos apenas pela própria sequência conversacional. Também um discurso humano está submetido a essa exigência. A diferença educacional que nos importa surge quando alguém apresenta uma conclusão sem conseguir participar do exame das razões que a sustentariam.

Não seria correto dizer que a IAG não tem ninguém a quem recorrer e, por isso, desaparece toda responsabilidade. Há instituições, desenvolvedores, provedores e usuários com atribuições diferentes. Miao e Holmes (2024, p. 23–25, 30) recomendam preservar a agência humana e não transferir aos sistemas a responsabilidade pelas decisões. A ausência de um "pai"

no sentido simples da autoria individual de cada frase não significa ausência de responsáveis. Ela exige tornar visíveis as responsabilidades distribuídas, sem concentrar todos os encargos no aluno nem absolver as organizações que oferecem e adotam a tecnologia.

4.3 O Deslocamento da Pergunta na Escola

O que a inteligência artificial generativa desfez foi a suficiência de um sinal, e não a necessidade de um critério. Isso desloca a pergunta que nos cabe fazer na escola. Não se trata, neste estudo, de determinar se a máquina pensa, questão que o procedimento adotado não permite resolver. Trata-se de saber se aquilo que ela devolve chega a inscrever-se em alguém. Essa é uma pergunta que não se responde examinando apenas a máquina.

A inscrição de que falamos, por aproximação interpretativa com o *Fedro*, não é a reprodução literal de um conteúdo. Designa a possibilidade de o aluno reconstruir um raciocínio, explicar a escolha de uma evidência, identificar uma dificuldade e rever uma conclusão. Nenhuma dessas operações funciona isoladamente como prova definitiva de aprendizagem. Reunidas em um acompanhamento, podem tornar visível um percurso que o produto final encobre. O interesse pedagógico recai no que o aluno consegue fazer com o discurso, e não somente em quem o produziu primeiro.

Uma investigação empírica ajuda a manter essa distinção. Bastani *et al.* (2025) realizaram um estudo randomizado com cerca de mil alunos de matemática em uma escola de ensino médio na Turquia. Os grupos com tutores generativos melhoraram o desempenho durante a prática assistida; na avaliação sem o recurso, o grupo com acesso menos orientado apresentou desempenho inferior ao controle. O tutor com salvaguardas pedagógicas mitigou esse prejuízo, sem produzir ganho autônomo significativo sobre o controle. O resultado não deve ser generalizado para toda disciplina ou uso de IA, mas mostra por que desempenho com apoio e aprendizagem sem o apoio precisam ser examinados separadamente.

O tempo de aprender reaparece, então, como condição de retorno e elaboração. Na sequência do *Fedro*, Sócrates contrapõe os jardins de Adônis, que florescem rapidamente, ao cultivo do agricultor que aguarda a maturação de suas sementes (Platão, 2016, p. 197–199, 276b–276e). A analogia permite distinguir o efeito vistoso de um resultado do cuidado com sua duração e fecundidade. Em educação, interessa perguntar o que continuará possível ao aluno depois que a resposta já não estiver diante dele.

Daí não decorre que aprender seja necessariamente lento em todos os momentos. Uma explicação oportuna pode dissipar rapidamente uma dificuldade. Tampouco decorre que toda ajuda deva ser retirada: muitos apoios ampliam o acesso e podem permanecer como parte legítima da atividade. O que precisa ser preservado é a oportunidade de o aluno participar das operações que constituem o objetivo da aprendizagem. O critério de autonomia deve ser definido em relação a esse objetivo, e não como ideal de independência absoluta de instrumentos e de outras pessoas.

O problema admite desdobramentos didáticos que não se reduzem à vigilância. Como proposta a investigar, uma atividade pode pedir ao aluno uma primeira tentativa, permitir posteriormente a consulta a um assistente e solicitar a explicação de uma mudança realizada. Em outra, duas respostas podem ser comparadas com uma fonte primária, cabendo ao aluno indicar onde uma interpretação se sustenta e onde excede a evidência. O material gerado deixa de ocupar o lugar da conclusão e passa a integrar o trabalho de leitura e julgamento.

Essas propostas não exigem que todo aluno demonstre saber em uma arguição oral improvisada. A elaboração pode tornar-se visível por versões comentadas, registros escritos, resolução de um novo caso, diálogo entre pares ou outras formas acessíveis de expressão. O tempo de responder, as condições de comunicação e os apoios necessários precisam ser considerados. Se a avaliação transformar hesitação ou dificuldade de expressão em prova de ausência de conhecimento, terá apenas substituído um indício frágil por outro.

O argumento possui limites. Apoiase em uma leitura delimitada de Platão, não em uma reconstrução integral de sua filosofia. A seleção bibliográfica é intencional e não autoriza afirmações sobre toda a recepção do *Fedro*. Os sistemas generativos mudam, e as evidências citadas dizem respeito às versões e condições examinadas. Além disso, a passagem do discurso inscrito na alma para a apropriação educacional contemporânea é uma interpretação assumida, não uma identidade histórica. Esses limites não invalidam a pergunta, mas determinam o alcance da resposta que podemos oferecer.

Pesquisas posteriores poderão comparar sequências didáticas nas quais a IAG fornece soluções completas, pistas graduais ou objeções às tentativas dos alunos. Interessa acompanhar não apenas a qualidade do trabalho entregue, mas a capacidade de justificar revisões, reconhecer limites e utilizar o que foi aprendido em novas situações. O problema filosófico orienta o que observar; não predetermina o resultado. Uma investigação assim poderá inclusive mostrar usos formativos que uma condenação geral impediria de reconhecer.

5 CONSIDERAÇÕES FINAIS

O retorno ao *Fedro* mostrou que a crítica à escrita contém movimentos diferentes. Thamous distingue a promessa do inventor dos efeitos que atribui à invenção e separa memória, recordação e aparência de sabedoria. Sócrates, na sequência, examina a invariância do escrito, sua circulação sem discernimento do destinatário e sua necessidade de assistência. A distinção entre essas objeções foi necessária para evitar que a comparação com a IAG dependesse de um critério único e insuficiente.

O assistente conversacional modifica uma condição evidente do escrito fixo: pode produzir novas réplicas. A mudança deve ser reconhecida, porque uma crítica que a negasse começaria por descrever mal seu objeto. Mas a variabilidade da resposta não demonstra, sozinha, a validade de suas razões, a pertinência da intervenção ou a aprendizagem de quem a recebe. Essas exigências permanecem abertas ao exame, mesmo quando o diálogo parece prosseguir sem dificuldade.

Também foi necessário deslocar a noção de defesa. Nem toda concessão é fraqueza, nem toda resistência é conhecimento. Uma resposta se sustenta por razões que possam ser examinadas e revistas, e não pela capacidade de repetir sua conclusão ou produzir uma sequência convincente de justificativas. A origem automática não torna um argumento necessariamente falso; sua fluência não o torna verdadeiro. O trabalho crítico começa quando nenhuma dessas características é aceita como substituta da análise.

Na escola, a consequência é concreta: o produto apresentado não pode absorver toda a atenção que deveria ser dirigida ao processo de aprender. A pergunta pela participação da IA é relevante, mas não basta. Importa acompanhar o que o aluno consegue compreender, reconstruir, contestar e realizar com os recursos de que dispõe. Isso exige situações de elaboração, critérios públicos e formas de expressão que não confundam rapidez, segurança verbal e conhecimento.

A resposta que não se defende é, nesse sentido, a resposta cuja apresentação pretende dispensar o trabalho de examinar e sustentar suas razões. Pode ter sido gerada por uma máquina ou repetida por uma pessoa. O risco educacional aparece quando sua disponibilidade ocupa o lugar da aprendizagem e quando a facilidade de produzir a próxima formulação adia indefinidamente o encontro com aquilo que ainda não foi compreendido.

Theuth e Thamous não decidem por nós o que fazer com a inteligência artificial generativa. Seu desacordo nos devolve a tarefa de julgar uma técnica a partir do que ela

possibilita e do que pode tornar dispensável em quem aprende. A escola não precisa competir com a velocidade da resposta. Precisa preservar a situação em que alguém possa deter-se nela, interrogá-la e, por fim, dizer por que a aceita, por que a modifica ou por que ainda não sabe.

REFERÊNCIAS

AGUIAR, Carlos Eduardo Souza. Farmacologia e filosofia de tecnologia: a contribuição do diálogo Fedro de Platão. **O que nos faz pensar**, Rio de Janeiro, v. 31, n. 53, p. 178–194, jul./dez. 2023. DOI: <https://doi.org/10.32334/oqnf.2023n53a940>. Disponível em: <https://oquenofazpensar.fil.puc-rio.br/oqnf/article/view/940>

BACK, Rainri. Compreensão e aprendizagem: uma abordagem hermenêutica. **Educação e Pesquisa**, São Paulo, v. 49, e256358, 2023. DOI: <https://doi.org/10.1590/S1678-4634202349256358>. Disponível em: <https://www.scielo.br/j/ep/a/MJ7vdgrvR6Njz8fQDy4PcTK/>

BAI, Yuntao et al. Constitutional AI: harmlessness from AI feedback. [S. l.]: **arXiv**, 2022. Preprint. DOI: <https://doi.org/10.48550/arXiv.2212.08073>. Disponível em: <https://arxiv.org/abs/2212.08073>

BASTANI, Hamsa et al. Generative AI without guardrails can harm learning: evidence from high school mathematics. **Proceedings of the National Academy of Sciences**, [S. l.], v. 122, n. 26, e2422633122, 2025. DOI: <https://doi.org/10.1073/pnas.2422633122>. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12232635/>

LARROSA, Jorge. Notas sobre a experiência e o saber de experiência. Tradução de João Wanderley Geraldi. **Revista Brasileira de Educação**, Rio de Janeiro, n. 19, p. 20–28, jan./abr. 2002. DOI: <https://doi.org/10.1590/S1413-24782002000100003>. Disponível em: <https://www.scielo.br/j/rbedu/a/Ycc5QDzZKcYVspCNspZVDxC/>

MIAO, Fengchun; HOLMES, Wayne. **Guia para a IA generativa na educação e na pesquisa**. Tradução de Teresa Margarida Loureiro Cardoso e Viviane Cristina Marques. Paris: UNESCO, 2024. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000390241>

PLATÃO. **Fedro**. Tradução e apresentação de José Cavalcante de Souza. Posfácio e notas de José Trindade Santos. São Paulo: Editora 34, 2016. Edição bilíngue.

RODRIGUES, Reginaldo Ferreira. Escrita e memória no Fedro de Platão. **Griot: Revista de Filosofia**, Amargosa, v. 11, n. 1, p. 100–115, jun. 2015. DOI: <https://doi.org/10.31977/grifi.v11i1.640>. Disponível em: <https://periodicos.ufrb.edu.br/index.php/griot/article/view/640>

SHARMA, Mrinank et al. Towards understanding sycophancy in language models. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 12., 2024, Vienna. **Proceedings [...]**. [S. l.]: ICLR, 2024. Disponível em: https://proceedings.iclr.cc/paper_files/paper/2024/file/0105f7972202cid4fb817da9f21a9663-Paper-Conference.pdf.

