

Estudo do MIT: Usar o ChatGPT não o tornará burro (a menos que você o faça de forma errada)

Um estudo sutil de IA, que você vai gostar!

[Alberto Romero](#)

MIT Study: Using ChatGPT Won't Make You Dumb (Unless You Do It Wrong)

THE ALGORITHMIC BRIDGE *by Alberto Romero*

O MIT publicou um [estudo](#) sobre como o uso de modelos de linguagem grandes (LLMs), como o ChatGPT, afeta seu cérebro e sua capacidade cognitiva. São 200 páginas - eu já li, então você não precisa ler.

As descobertas são relevantes e surpreendentemente matizadas. A crítica dos autores aos LLMs - porque este é um artigo que critica o abuso de LLMs em ambientes educacionais - é apoiada e qualificada pelos resultados - não há preto ou branco aqui.

O aspecto mais importante do estudo é que ele usa um método de monitoramento cerebral (eletroencefalografia, EEG) que mede diretamente a atividade cerebral. Isso é fundamental porque não depende de dados comportamentais ou de autorrelato (o que as pessoas dizem que pensam ou sentem), que geralmente têm validade duvidosa.

Este é o seu cérebro lhe dizendo o que acontece quando você abusa do ChatGPT.

I. Projeto experimental: redação

Para lhe dar uma imagem mental do estudo, vou começar com o projeto experimental.

Os autores reuniram 54 pessoas e as colocaram em três grupos: Grupo LLM, grupo de mecanismos de busca (vou deixá-los de lado) e grupo do cérebro. Cada participante teve permissão para usar os materiais atribuídos ao seu grupo, o que significa que o grupo do cérebro não teve acesso a nada além de seus conhecimentos e habilidades.

O estudo foi dividido em quatro sessões de 20 minutos cada, ao longo de quatro meses. Nas três primeiras sessões, cada participante teve de escrever uma redação com base em uma sugestão como esta (dos SATs):

Muitas pessoas acreditam que lealdade, seja a um indivíduo, a uma organização ou a uma nação, significa apoio incondicional e inquestionável, não importa o que aconteça. Para essas pessoas, a retirada do apoio é, por definição, uma traição à lealdade. Mas a verdadeira lealdade não exige, às vezes, que critiquemos aqueles a quem somos leais? Se percebermos que eles estão fazendo algo que acreditamos estar errado, a verdadeira lealdade não exige que nos manifestemos, mesmo que tenhamos de ser críticos?

Na quarta sessão opcional - a mais importante -, os autores trocaram os grupos de LLM e cérebro: aqueles que haviam usado um LLM para escrever durante as três sessões anteriores agora tinham de usar apenas o cérebro, e vice-versa, aqueles que usaram apenas o cérebro agora tinham de usar um LLM. 18 pessoas voltaram para a quarta sessão e foram instruídas a escolher um tópico com o qual haviam se familiarizado em qualquer uma das sessões anteriores. Esses grupos são chamados, respectivamente, de LLM-para-cérebro e cérebro-para-LLM.

Aqui está um diagrama visual do estudo:

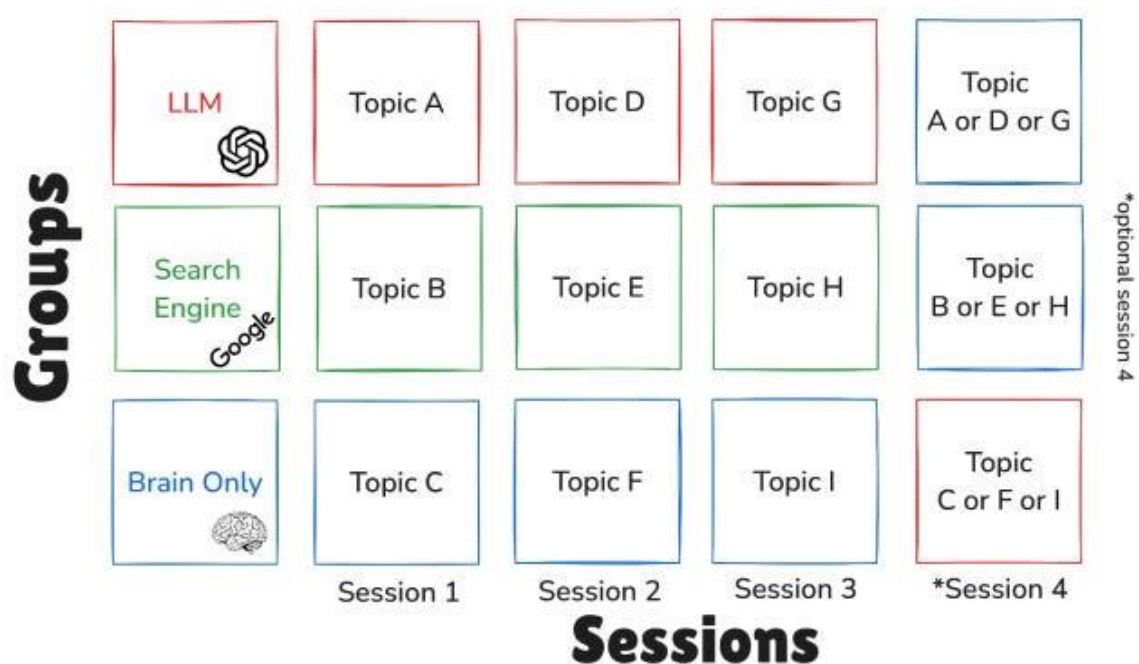


Diagrama visual do estudo. Fonte: [MIT](#)

II. Cinco limitações para pesquisas futuras

Antes de analisar os resultados, aqui estão cinco limitações deste estudo (os autores reconhecem que há mais trabalho a ser feito com relação a esses aspectos):

1. A amostra é pequena, o que significa que o estudo pode não ter o poder estatístico necessário para detectar de forma confiável efeitos pequenos ou moderados; para isso, seria necessária uma amostra muito maior. A amostra também é homogênea: as pessoas nas proximidades do MIT certamente não refletem a distribuição das pessoas em todo o mundo: todas têm ensino superior, alto nível socioeconômico, são americanas, jovens etc.
2. Eles usaram o ChatGPT com o GPT-4o. Eles mencionam que não se deve extrapolar os resultados para outros LLMs, mas acho que estão sendo excessivamente cautelosos. Não creio que haja qualquer razão para que isso não se aplique a qualquer outro LLM (incluindo modelos de raciocínio como o o3). De qualquer forma, mais pesquisas nesse sentido seriam bem-vindas.
3. Eles se concentraram em uma tarefa de redação em um ambiente educacional. Esses resultados podem não se aplicar a atividades de trabalho ou a ambientes da vida real e podem não se manter em todas as tarefas (escrever uma redação é altamente específico). Eles também não dividiram "escrever uma redação" em subtarefas para analisar se os LLMs prejudicam o desempenho em um subconjunto delas e ajudam em um subconjunto diferente.
4. Eles usaram EEG, que permite medir a atividade cerebral, mas esse método tem baixa resolução espacial, o que significa que é difícil identificar exatamente onde as coisas estão acontecendo. A resolução temporal é alta porque os sinais bioelétricos viajam rapidamente. Um estudo que usasse um método de neuroimagem mais potente, como a fMRI, com grande resolução espacial (pior resolução temporal), poderia fornecer informações importantes.
5. Estudo transversal: Os autores observaram três grupos em quatro sessões, mas não realizaram um acompanhamento mais longo para avaliar o impacto a longo prazo. Um estudo longitudinal (medido por longos períodos) forneceria informações adequadas para saber se o efeito do uso de LLMs é temporário ou permanente.

III. Análise linguística e de EEG: Grupo somente LLM

Os autores dividiram sua análise na parte de processamento de linguagem natural (vou chamá-la de análise linguística para simplificar) e na parte de EEG. Embora o EEG seja indiscutivelmente mais revelador, achei a análise linguística bastante perspicaz, portanto, compartilharei ambas. Aqui está o que eles descobriram sobre o grupo LLM. Grupo do cérebro na próxima seção.

Análise linguística

1. Nas sessões de 1 a 3, os participantes melhoraram progressivamente na escrita de prompts variados, mas, por fim (sessão 3), abordaram a tarefa de maneira "copiar-colar com pouco esforço", resultando em redações que poderiam ser rotuladas como "redações do tipo ChatGPT". Os participantes não se preocuparam em distanciar suas redações do que o ChatGPT escreveria por padrão. (Observação: é difícil discernir o quanto a variação entre as sessões 1, 2 e 3 responde a uma redução do envolvimento cognitivo induzida pelo LLM em comparação com o cansaço).

2. Baixa percepção de propriedade e baixa capacidade de citar a redação: O ChatGPT co-escreveu o ensaio, o que fez com que os participantes se sentissem distantes. Isso é bom. A alternativa é sentir como seu algo que não é. Por outro lado, a incapacidade de citar algo que teve seu nome associado a ele é algo ruim. Isso implica memória superficial e codificação de conteúdo semântico, o que se traduz em "você não se lembra ou não entende o que escreveu em conjunto".
3. Os juízes humanos notaram pouca distância entre as redações dos participantes (eles não sabiam quem as escreveu a priori). Isso significa, como esperado, que as pessoas que usaram o ChatGPT estavam escrevendo *praticamente o mesmo material* em termos de estrutura e abordagem. Isso está de acordo com a ideia generalizada de que os LLMs levam você ao centro da distribuição de palavras, mas também do significado, estilo, estrutura etc.
4. Redações mais longas: O grupo de LLMs que escreve ensaios mais longos é trivial porque os LLMs são muito mais rápidos do que os humanos para escrever, são excessivamente prolixos por padrão e nunca se cansam. 20 minutos para eles escreverem algo sobre lealdade ou arte é uma eternidade para nós. (Os humanos ainda tinham que fazer a integração, mas a compilação é mais fácil do que a criação).

Análise de EEG

1. Os participantes demonstraram uma diminuição do envolvimento cognitivo ao longo do tempo devido a uma configuração familiar, o que significa que "você não desperdiça recursos mentais se não precisar", o que torna o uso da IA atraente do ponto de vista do "orçamento cognitivo".
2. No entanto, há consequências para isso: Os autores encontraram padrões de conectividade (quais partes do cérebro você está usando) reduzidos em cerca de 50% e também reduziram a ativação neural em geral (o quanto você está usando) em comparação com o grupo do cérebro. Isso significa que o grupo LLM envolveu menos seus cérebros em geral. Os autores chamam isso de "modo cognitivo automatizado e estruturado".
3. Na sessão 2, há um pico de atividade de "fluxo de integração", o que significa que os usuários de LLM (também usuários de mecanismos de pesquisa) se envolvem na atividade de integrar grandes quantidades de informações que lhes foram entregues. É como pedir a um estagiário que organize um monte de documentos sem ter que encontrá-los e selecioná-los previamente ou entendê-los depois.

Há um detalhe crítico nos limites de tempo do experimento que quero mencionar aqui como um aparte.

O fato de ter 20 minutos para escrever uma redação implicitamente leva os participantes do grupo LLM a se concentrarem em copiar e colar em vez de incorporar seus pensamentos e desafiar o ChatGPT. Poderíamos pensar que, com mais tempo, as pessoas não usariam os LLMs sem pensar. Entretanto, minha hipótese é que esse

projeto experimental imita bem as restrições naturais da vida real. 20 minutos é um bom substituto para: "O prazo é amanhã", ou "Tenho que lavar a roupa", ou "Preferia estar jogando videogame".

Esse "modo cognitivo automatizado e estruturado" é como as pessoas usam naturalmente o ChatGPT quando estão limitadas pelos deveres e responsabilidades da vida real. O ChatGPT atua, antes de mais nada, como um *descarregador cognitivo*.

IV. Análise linguística e de EEG: Grupo somente com cérebro

Análise linguística

1. As redações do grupo do cérebro foram mais curtas em geral (naturalmente) e também mais diversificadas. Isso se aplica às sessões 1, 2 e 3. A propriedade percebida e a capacidade de citar permanecem altas em todo o estudo para esse grupo.
2. Isso significa que os participantes que usam o cérebro se inspiram em seu conhecimento, suas opiniões, suas experiências e sua identidade de uma forma que as pessoas que usam LLMs não fazem (pelo menos sob restrições críticas de tempo). O cérebro também está disponível para esse último grupo, mas o ChatGPT, atuando como um descarregador cognitivo, impede que eles aproveitem a rica fonte que é a vida deles e, em vez disso, aproveitam a fonte homogênea e sem alma que é a Internet compactada.

Análise de EEG

1. O cérebro mostra alta atividade neural constante (intensidade do envolvimento) e padrões de conectividade (escopo do envolvimento). O cérebro permanece "envolvido na integração semântica, na ideação criativa e no automonitoramento executivo". E eles acrescentam: "O grupo somente com cérebro aproveitou redes neurais amplas e distribuídas para conteúdo gerado internamente". Se você tiver apenas o cérebro à sua disposição, terá que fazer o trabalho duro por conta própria. Você mantém suas ferramentas nativas afiadas.

V. Sessão 4: LLM-para-cérebro e cérebro-para-LLM

18 participantes participaram da quarta sessão. Esta parte é a mais relevante para aqueles que desejam abordar a IA de forma prática e útil, mas sem prejudicar sua capacidade cognitiva.

Lembre-se: LLM-para-cérebro são pessoas que usaram LLMs durante as três primeiras sessões e agora devem usar apenas o cérebro. Do cérebro para o LLM, o contrário. Eles foram instruídos a escrever uma redação sobre um tópico com o qual já estivessem familiarizados nas sessões anteriores (isso é fundamental para remover fatores de confusão).

LLM para o cérebro

1. Eles "usaram N-gramas de sessões anteriores de LLM". Isso significa que, por ter usado LLMs antes, o cérebro influencia seu vocabulário e expressões para que se assemelhem aos de um LLM. Essa é uma péssima notícia para as pessoas que usam o ChatGPT para aprender a escrever. De certa forma, elas estão se tornando *papagaios estocásticos de um papagaio estocástico*.
2. Uma descoberta surpreendente aqui: Eles "obtiveram pontuação mais alta pelos professores humanos dentro do grupo". Dentro do grupo significa que essas pessoas, que vieram do uso de LLMs por uma hora, obtiveram uma pontuação melhor aos olhos dos professores do que aquelas que usaram seus cérebros sozinhas nas três sessões anteriores. Isso é fundamental para o sistema educacional. Isso significa que o cérebro de um escritor não profissional pode não ser a melhor ferramenta de escrita para ele, portanto, ele é incentivado a usar o ChatGPT. Isso aumenta sua dependência de uma ferramenta externa, obtendo ganhos substanciais a curto prazo (notas mais altas) às custas de uma deterioração cognitiva a longo prazo (menor habilidade mental).
3. Os autores colocam a questão da seguinte forma: "Os participantes originais do LLM podem ter ganhado na aquisição inicial de habilidades usando o LLM para uma tarefa, mas isso não substituiu a integração neural mais profunda, que pode ser observada no grupo original somente com o cérebro. As intervenções educacionais devem considerar a combinação de assistência de ferramentas de IA com fases de aprendizado sem ferramentas para otimizar a transferência imediata de habilidades e o desenvolvimento neural de longo prazo."

Brain-to-LLM

1. Isso é ótimo: "Melhor integração de conteúdo em comparação com as sessões anteriores do Brain. . . . Mais solicitações de busca de informações. Pontuação, em sua maioria, acima da média em todos os grupos." Isso significa que usar o ChatGPT após ter escrito uma redação e ter se envolvido com ela - os participantes não sabiam do que se tratava a quarta sessão, portanto, seu envolvimento nas três sessões anteriores foi sério e honesto - *é melhor* do que usar apenas o cérebro. A integração é melhor, os prompts são melhores (envolvimento com a tarefa > ter experiência anterior em LLM) e a qualidade é melhor. (Observe que isso se aplica a escritores não profissionais, que se beneficiariam muito menos com o uso do ChatGPT).
2. A IA não é prejudicial, *mas útil* se você envolver seu cérebro com frequência e intensidade suficientes: "O reengajamento apoiado pela IA invocou altos níveis de integração cognitiva, reativação da memória e controle de cima para baixo. Por outro lado, o uso repetido do LLM nas sessões 1, 2 e 3 do grupo LLM original refletiu a redução da conectividade ao longo do tempo. Esses resultados enfatizam a interação dinâmica entre a estrutura cognitiva e o envolvimento neural em contextos de aprendizagem com suporte de IA."
3. De certa forma, essas descobertas reforçam minha tese da [regra de ouro do uso da IA](#): "A IA é uma ferramenta e, como qualquer outra, deve seguir a regra de ouro: Todas as ferramentas devem aprimorar, nunca corroer, a mais importante delas: sua mente. Tenha curiosidade sobre a IA, mas também

examine como ela molda seus hábitos e seus padrões de pensamento. Siga essa regra e você não terá nada a temer." Dizer que "a IA é ruim" é simplesmente errado. O contexto, o escopo, o esforço pessoal etc. são importantes ao avaliar o valor do uso de ferramentas de IA.

VI. Conclusões gerais

Como eu disse, esse estudo não é preto ou branco (independentemente do que os influenciadores digam). Há muitas descobertas interessantes, inclusive como usar o LLM sozinho é prejudicial, como usar o cérebro é bom e como combiná-los *na ordem certa e na medida certa* pode ser uma ótima opção. Aqui estão algumas conclusões gerais adicionais selecionadas do artigo:

1. **Brain-to-LLM >>> LLM-to-Brain:** "Os participantes do Brain-to-LLM apresentaram maior conectividade neural do que os do grupo LLM nas sessões 1, 2 e 3. Isso sugere que reescrever uma redação usando ferramentas de IA (depois de uma redação anterior sem IA) envolveu interações mais extensas da rede cerebral. Em contraste, o grupo LLM-para-Cérebro, exposto ao uso prévio de LLM, demonstrou um esforço neural menos coordenado na maioria das faixas, bem como um viés no vocabulário específico de LLM."
2. **O uso de IA altera o envolvimento do cérebro e não apenas o desempenho da tarefa:** "Coletivamente, essas descobertas apoiam a visão de que as ferramentas de suporte externo reestruturam não apenas o desempenho da tarefa, mas também a arquitetura cognitiva subjacente. O grupo Brain-only aproveitou redes neurais amplas e distribuídas para conteúdo gerado internamente; . . . e o grupo LLM otimizou a integração processual de sugestões geradas por IA."
3. **Se você confiar muito na IA, ficará mais burro:** "As ferramentas de IA, embora valiosas para apoiar o desempenho, podem prejudicar involuntariamente o processamento cognitivo profundo, a retenção e o envolvimento autêntico com o material escrito. Se os usuários dependerem muito das ferramentas de IA, eles poderão alcançar uma fluência superficial, mas não conseguirão internalizar o conhecimento ou ter um senso de propriedade sobre ele."
4. **Pode haver uma sequência ideal de uso da ferramenta:** "Do ponto de vista educacional, esses resultados [sessão 4] sugerem que o momento estratégico da introdução da ferramenta de IA após o esforço inicial autodirigido pode aumentar o envolvimento e a integração neural. Os marcadores de EEG correspondentes indicam que essa pode ser uma sequência mais ideal do ponto de vista neurocognitivo do que o uso consistente de ferramentas de IA desde o início."
5. **Você deve evitar incorrer em dívida cognitiva:** (Conclusão preliminar, mais pesquisas são necessárias.) "Quando as pessoas não conseguem se envolver criticamente com um assunto, sua redação pode se tornar tendenciosa e superficial. Esse padrão reflete o acúmulo de dívida cognitiva, uma condição na qual a dependência repetida de sistemas externos, como os LLMs, substitui os processos cognitivos de esforço necessários para o pensamento independente. A dívida cognitiva adia o esforço mental no curto prazo, mas resulta em custos

de longo prazo, como a diminuição da investigação crítica, o aumento da vulnerabilidade à manipulação e a diminuição da criatividade."

Isso é tudo! Foi muito divertido. Espero que você tenha achado útil. Eu certamente achei útil escrevê-lo (não usei o ChatGPT para descarregar a carga cognitiva).

Ficarei atento a quaisquer estudos como esse que eu encontrar, especialmente se eles incluírem métodos de neuroimagem (fMRI) ou se forem longitudinais (teremos que esperar por isso).