

Perspectiva filosófica, Recife, v. 52 n. 3 (edição especial) (2025): Filósofas Analíticas Brasileiras. Universidade Federal de Pernambuco. ISSN 0104-6454. DOI: <https://doi.org/10.51359/2357-9986.2025.265938>

Submetido em: 1 mar. 2025
Aceito em: 18 ago. 2025

Podem argumentos mudar opiniões?

*Can arguments change minds?*¹

Catarina Dutilh Novaes²

Vrije Universiteit Amsterdam (VUA)

Ariana Romão dos Reis (trad.)³

Universidade Federal do Maranhão (UFMA)

Ederson Saфра Melo (trad.)⁴

Universidade Federal do Maranhão (UFMA)

RESUMO

Argumentos podem mudar opiniões? Filósofos gostam de pensar que sim. Contudo, uma vasta quantidade de evidências empíricas sugere que argumentos não são ferramentas muito eficientes para mudar opiniões. Como lidar com as diferentes avaliações do potencial dos argumentos para mudar opiniões? Para abordar essa questão, devemos levar em conta os contextos mais abrangentes em que os argumentos ocorrem, em particular a propagação de mensagens por meio de redes de atenção e as escolhas que agentes epistêmicos devem fazer entre potenciais fontes alternativas de conteúdo e informação, que são muito influenciadas pelas percepções de credibilidade e confiabilidade. Argumentos podem mudar opiniões, mas apenas sob condições socioepistêmicas propícias e favoráveis.

Palavras-chave: mudança de opinião. troca epistêmica. argumentação. bolha epistêmica. câmara de eco.

ABSTRACT

Can arguments change minds? Philosophers like to think that they can. However, a wealth of empirical evidence suggests that arguments are not very efficient tools to change minds. What to make of the different assessments of the mind-changing potential of arguments? To address this issue, we must take into account the broader contexts in which arguments occur, in particular the propagation of messages

¹ Dutilh Novaes, C. VII—Can Arguments Change Minds?, *Proceedings of the Aristotelian Society* 123 (2), julho 2023, pp.173-198. DOI 10.1093/arisoc/aoad006.

² E-mail: c.dutilhnovaes@vu.nl. ORCID: <https://orcid.org/0000-0002-7347-3795>.

³ E-mail: arianardr@gmail.com. ORCID: <https://orcid.org/0009-0006-8850-0561>.

⁴ E-mail: ederson.saфра@ufma.br. ORCID: <https://orcid.org/0000-0003-1706-1128>.

across net-works of attention, and the choices that epistemic agents must make between alternative potential sources of content and information, which are very much influenced by perceptions of reliability and trust-worthiness. Arguments can change minds, but only under conducive, favourable socio-epistemic conditions.

Keywords: mind-changing. epistemic exchange. argumentation. epistemic bubble. echo chamber.

INTRODUÇÃO

Argumentos podem mudar opiniões? Filósofos gostam de pensar que sim: ao engajarmo-nos no processo (presumivelmente racional) de considerar cuidadosamente razões a favor ou contra uma dada posição ou visão, ajustamos nossas crenças de acordo com tais razões⁵. Conforme essa visão otimista, notoriamente defendida por John Stuart Mill em particular, não apenas mudamos nossas opiniões ao sermos expostos a argumentos convincentes (uma afirmação descritiva), mas também melhoramos nossa posição epistêmica geral pelas considerações cuidadosas de razões (uma afirmação avaliativa).

No entanto, uma grande quantidade de evidências empíricas e anedóticas parece sugerir que argumentos não são, de fato, ferramentas muito eficientes para mudar opiniões (Gordon-Smith, 2019; McIntyre, 2021). Por exemplo, o fenômeno bem documentado da polarização (Isenberg, 1986; Sunstein, 2002) sugere que, quando expostas a argumentos que apoiam posições diferentes de suas opiniões prévias, as pessoas frequentemente (embora talvez nem sempre) tornam-se ainda mais convencidas de suas opiniões prévias, ao invés de serem persuadidas por argumentos (Olsson, 2013). Frequentemente, debates argumentativos se assemelham mais a jogos em que os participantes buscam ‘marcar pontos’ (Cohen, 1995; Dutilh Novaes, 2021), em vez de se engajarem em uma análise meticulosa de diferentes visões com o propósito de aprimoramento epistêmico.

O que fazer com essas diferentes avaliações do potencial dos argumentos para mudar opiniões? Para abordar essa questão, parece que precisamos olhar para além do conteúdo e da qualidade dos argumentos por si só:⁶ devemos considerar também os contextos mais amplos nos quais os argumentos ocorrem, em particular

⁵ Neste artigo, falo de ‘mudança de opinião’ de maneira pouco rigorosa, mas o conceito também pode ser tratado mais sistematicamente. Existem diferentes teorias formais que oferecem uma explicação sobre o que significa mudar de opinião de alguém, como probabilidades bayesianas e várias teorias de revisão de crenças. Para nossos propósitos aqui, as diferenças entre elas são irrelevantes, já que todas tratam de como agentes atualizam suas crenças à luz de informações recebidas.

a propagação de mensagens por redes de comunicação e as escolhas que agentes epistêmicos devem fazer entre potenciais fontes alternativas de conteúdo e informação. Essas escolhas são muito influenciadas por percepções de credibilidade e confiabilidade, o que significa que a fonte do argumento pode ser ainda mais decisiva do que seu conteúdo ou qualidade quando se trata de quão persuasivo ele será para uma determinada pessoa. (Nesse aspecto, a argumentação se assemelharia mais ao testemunho do que se poderia imaginar, como já argumentei em outro trabalho: Dutilh Novaes, 2020b.) Em suma: os argumentos podem muito bem mudar opiniões, porém, somente sob condições socioepistêmicas propícias e favoráveis.

Neste artigo, utilizo um modelo de troca epistêmica de três níveis, que venho desenvolvendo ao longo dos últimos anos (Dutilh Novaes, 2020b), para (esperançosamente) compreender os mecanismos envolvidos nesses processos e as condições sob as quais argumentos podem mudar opiniões. Começo com uma avaliação ‘otimista’ do poder da argumentação para mudar opiniões, conforme a formulação de John Stuart Mill, e suas limitações, tal como discutida na literatura (ao menos como uma descrição precisa do fenômeno em questão). Em seguida, ofereço uma breve descrição do modelo de três níveis e de sua relevância para o tema em questão. Na seção IV, discuto dois exemplos da vida real de pessoas que tiveram avanços epistêmicos que envolveram (ao menos em certa medida) o engajamento com argumentos, mas somente em um contexto de condições socioepistêmicas favoráveis. Na seção V, esclareço algumas questões pendentes. Concluo, então, com algumas considerações finais.

A CONCEPÇÃO MILLIANA DE ARGUMENTAÇÃO E SUAS LIMITAÇÕES

Mill é um dos principais expoentes da visão de que situações argumentativas interpessoais, envolvendo pessoas que realmente discordam umas das outras, têm o potencial de mudar opiniões (sobretudo para melhor, acredita ele)⁷. Em *Sobre a liberdade* (Mill, [1859] 2011), ele nota que, quando nossas ideias são desafiadas por

⁶ Considero que a qualidade de um argumento está relacionada aos critérios familiares de qualidade argumentativa, como validade e correção. A qualidade do argumento também pode ser analisada de forma probabilística.

⁷ Eu mesma defendi essa visão (Dutilh Novaes 2020a), mas com a importante ressalva de que o efeito epistêmico benéfico da argumentação interpessoal ocorrerá apenas em um contexto de circunstâncias específicas que garantam uma troca de ideias em boa-fé (por exemplo, dentro de uma comunidade de matemáticos). Veja abaixo uma discussão sobre as circunstâncias em que trocas argumentativas levam de maneira confiável a um aprimoramento epistêmico.

aqueles que discordam de nós, somos forçados a avaliar criticamente nossas próprias crenças:

Ele [O homem] é capaz de retificar os seus erros através da discussão e da experiência. Não simplesmente pela experiência; tem de haver discussão, que mostre como há de a experiência ser interpretada. Opiniões e práticas erradas cedem gradualmente a fatos e argumentos: mas os fatos e argumentos têm de ser colocados perante a mente, de modo a que possam ter efeito sobre ela. (Mill, [1859] 2011, p. 47)

Esse processo é frequentemente descrito como uma *troca livre de ideias* e, de acordo com Mill, é benéfico mesmo quando estamos certos e nossos interlocutores estão errados. O resultado esperado é que as crenças remanescentes, aquelas que sobreviveram aos desafios críticos, estarão melhor justificadas do que as crenças mantidas antes desses debates. Como Mill coloca: "Assim que não há qualquer inimigo no terreno, tanto professores como aprendizes adormecem no posto" ([1859] 2011, p. 74). As divergências, assim, nos forçam a permanecer epistemicamente atentos, em vez de nos acomodarmos com crenças existentes e enraizadas — o que Mill descreve como ‘dogma morto’ (Simpson 2021).

Porém, para que esse processo seja bem-sucedido, os divergentes devem ter permissão para expressar livremente suas críticas e opiniões. De fato, a forte defesa da liberdade de expressão feita por Mill é uma de suas posições mais célebres. Um de seus principais argumentos a favor da liberdade de expressão é de natureza epistêmica: ele enfatiza o papel desempenhado pela livre troca de ideias em facilitar o crescimento do conhecimento em uma sociedade. Quanto mais visões divergentes e argumentos a favor ou contra cada uma delas forem trocados, maior será a probabilidade de que as ‘melhores’ visões prevaleçam (Halliday e McCabe, 2019).

No entanto, não basta que os divergentes tenham a oportunidade de expressar suas opiniões livremente; é também de crucial importância que os receptores dessas opiniões e argumentos estejam dispostos a engajar-se de boa-fé e com a mente aberta.⁸ Mill dá bastante atenção às condições estruturais para a livre troca de ideias (em particular, à necessidade de que não haja nenhum tipo de censura sancionada pelo Estado). Contudo, Mill parece não levar suficientemente em conta nossas tendências, bem documentadas, de evitar o engajamento com

⁸ Há também a importante questão (que será discutida brevemente) de se as vozes divergentes conseguirão atrair atenção, por exemplo, se pertencerem a grupos marginalizados.

visões divergentes ou de explicar evidências contrárias para preservarmos as nossas crenças prévias (um ponto que será discutido em breve).

Mais recentemente, Alvin Goldman articulou uma visão semelhante sobre a epistemologia social da argumentação (Goldman, 1994, 2004). O ponto de partida para Goldman é o reconhecimento de uma situação de divisão epistêmica do trabalho, onde diferentes membros de uma comunidade epistêmica possuem conhecimentos distintos e, assim, podem se beneficiar da troca desses diferentes recursos epistêmicos entre si. Além disso, dada a nossa inescapável falibilidade, essas trocas com outros conhecedores podem ajudar a expor nossas próprias crenças equivocadas (como também observado por Mill). Um terceiro aspecto de nossa situação socioepistêmica é que, às vezes, as pessoas possuem incentivos para enganar e induzir outras pessoas ao erro, tornando necessária uma certa dose de vigilância epistêmica. Nesse contexto, é contra essas condições que, segundo Goldman, a argumentação se torna uma ferramenta valiosa na busca pela verdade e na prevenção de erro:

As normas de boa argumentação são substancialmente dedicadas à promoção do discurso verídico e à exposição da falsidade, seja ela intencional ou não. ... Normas de boa argumentação fazem parte de uma prática que incentiva a troca de verdades por meio de discursos sinceros, diligentes e mutuamente corretivos. (Goldman, 1994, p. 30)

Mas será que a argumentação consegue *confiavelmente* promover a verdade e evitar o erro em contextos socioepistêmicos, conforme sugerido por Mill e Goldman? Será que revisamos prontamente nossas crenças quando expostos a (bons) argumentos que as contradizem? Será que de fato “gradualmente cedemos aos fatos e argumentos”, como afirma Mill? Parece que Mill e Goldman são excessivamente otimistas quanto ao poder dos argumentos de mudar opiniões. Mas, na verdade, a argumentação parece ser uma maneira um tanto ineficiente para mudar opiniões em muitas situações da vida real (Gordon-Smith, 2019).

De fato, parece que as pessoas geralmente evitam revisar suas visões sobre crenças firmemente enraizadas (um ponto notoriamente destacado por Quine (1951). Quando confrontadas com argumentos ou evidências que contradizem suas crenças, as pessoas tendem ou a ignorar as evidências, ou explicá-las — como sabemos da literatura sobre viés de confirmação (Nickerson, 1998) —, ou a

descreditar a fonte do argumento como não confiável.⁹ Essas tendências são exemplificadas pelos chamados negacionistas da ciência, tais como os terraplanistas (McIntyre, 2021), mas também na prática científica, onde paradigmas consolidados frequentemente resistem a uma quantidade significativa de evidências contrárias antes que ocorra uma ‘revolução científica’ (Kuhn, 1962). Em particular, argumentos que ameaçam crenças centrais, sentimentos de pertencimento e identidades (por exemplo, crenças políticas) parecem desencadear várias formas de raciocínio motivado, nas quais se ignora ou rejeita aqueles argumentos sem um engajamento substancial com seu conteúdo (Taber e Lodge, 2006; Kahan, 2017). Engajar-se (ou não) na argumentação é frequentemente um meio de expressar e consolidar identidades sociais, em vez de se aproximar da verdade (Talisso, 2019; Hannon, 2019).

Além disso, ao escolher entre uma vasta gama de opções, há uma tendência a gravitar em direção a conteúdos e fontes que confirmem as opiniões já existentes, nas chamadas ‘câmaras de eco’ e ‘bolhas epistêmicas’ (Nguyen, 2020). Conversas com pessoas de opiniões semelhantes podem reforçar crenças prévias e até mesmo conduzir pessoas a versões mais extremas dessas crenças (Olsson, 2013). Isso significa que a mera disponibilidade de opiniões divergentes não é suficiente para garantir que os conhecedores permaneçam epistemicamente atentos e considerem todos os lados de uma questão. Sempre existe a opção de ignorar (isto é, de não engajar-se com) esses divergentes e o conteúdo de seus argumentos, especialmente se forem percebidos como não confiáveis (Dutilh Novaes, 2020b). Este é o fenômeno familiar da polarização: em vez de aproximar as partes, a argumentação e a deliberação podem ter o efeito oposto, o efeito de afastar ainda mais as partes (Sunstein, 2002).

Outro obstáculo é o fato de que a ausência de censura sancionada pelo governo (como proposta por Mill) não garante que todas as vozes relevantes sejam verdadeiramente *ouvidas*. Visões divergentes defendidas por grupos sociais marginalizados tenderão a atrair menos atenção do que aquelas com proponentes poderosos; a chamada livre troca de ideias é aquela em que as diferenças de poder afetam significativamente a disseminação e a aceitação das visões. Esse é o conhecido problema da *inclusão* em sociedades democráticas (Young, 2000), que

⁹ Veja, contudo, Mercier (2020) e Coppock (2022), que argumentam que agentes epistêmicos regularmente, e competentemente, atualizam suas crenças diante de novas informações, inclusive em questões carregadas de valores, como a política.

tem consequências tanto políticas quanto epistêmicas. Na maioria das vezes, não é apenas a força ou qualidade de um argumento que determina sua aceitação; a posição social de seus proponentes é um fator decisivo para o quanto o argumento será disseminado e considerado persuasivo.

Certamente, *há* alguns contextos em que a troca de razões em interações argumentativas parece levar de maneira confiável as pessoas a mudarem de opinião e a obterem aprimoramento epistêmico (Mercier, 2018; Dutilh Novaes, 2020a, caps. 8 e 9).¹⁰ A literatura sobre resolução de problemas em grupo tem estabelecido que, no caso de problemas chamados de ‘intelectivos’ — ou seja, aqueles que possuem uma resposta única dentro de um dado arcabouço teórico, por exemplo, um problema matemático ou lógico — a discussão em grupo entre pares tem um benefício claro, o efeito de conduzir à verdade (Laughlin, 2011). De fato, em contextos especializados, como na ciência ou na matemática, o ‘embate’ argumentativo é uma forma essencial de produzir conhecimento (Longino, 1990; Lakatos, 1976). No entanto, isso é menos evidente no caso dos chamados ‘problemas de julgamento’, que dizem respeito a valores e juízos para os quais não há uma única resposta correta (Laughlin, 2011). (A distinção entre problemas intelectivos e de julgamento é uma questão de grau.) É importante notar que, em situações da vida real, muito frequentemente nos deparamos com problemas de julgamento, e para esse tipo de problema, a literatura sobre resolução de problemas em grupo sugere que a argumentação não leva de maneira confiável a melhores resultados (isso não é surpreendente, dado que, em problemas de julgamento, muitas vezes não há consenso sobre o que constitui evidência conclusiva). De fato, muitos desses problemas são instâncias de *desacordos profundos* (Fogelin, 1985), que podem não ser passíveis de serem resolvidos por meio de raciocínio e argumentação.

Essas observações sugerem que não somos ‘propriamente millianos’ quando se trata de argumentação e de divergência. A atenção epistêmica que Mill acreditava que seria a consequência natural, quase automática, da exposição a opiniões e argumentos divergentes frequentemente não acontece. Assim, a

¹⁰ O conceito de ‘aprimoramento epistêmico’ pressupõe a existência de métricas adequadas que nos permitam medir o progresso. Uma métrica natural é simplesmente o que os epistemólogos chamam de acurácia, que grosseiramente corresponde à “busca pela verdade e prevenção do erro” de Goldman (veritismo). Contudo, métricas mais refinadas podem ser consideradas, por exemplo, como aprimoramento epistêmico em termos de *entendimento* (Grimm, Baumberger e Ammon 2016).

concepção milliana é descritivamente imprecisa ou, no mínimo, incompleta. Pode-se argumentar que a concepção milliana ainda é *normativamente* correta; contudo, dado que parece ser altamente idealizada, é discutível que seja adequada para oferecer recomendações *prescritivas* (no sentido de Bell, Raiffa e Tversky, 1988) para agentes humanos concretos.¹¹ Em vez disso, precisamos de uma abordagem mais realista para a epistemologia (social) da argumentação, uma que leve em consideração não apenas as limitações cognitivas dos conhecedores individualmente, mas também as complexidades sociais desses processos.

O MODELO DE TROCA EPISTÊMICA DE TRÊS NÍVEIS

Acabamos de ver que a troca livre de ideias é prejudicada por vários fatores, como relações estruturais de poder e tendências cognitivas e sociais, tanto que não há garantia de que opiniões e práticas erradas “cederão gradualmente aos fatos e argumentos”. Para abordar algumas das limitações da concepção milliana de argumentação, tenho desenvolvido um modelo de troca epistêmica de três níveis, que apresenta uma concepção mais realista de troca epistêmica por meio da argumentação, considerando os custos, obstáculos e riscos de engajar-se em trocas argumentativas (Dutilh Novaes, 2020b). Embora seja um modelo de processos epistêmicos sociais em geral, a ideia central é que a argumentação consiste realmente em uma *troca*, onde os recursos fluem em ambas as direções (do argumentador para o receptor, mas também do receptor para o argumentador), sendo, portanto, um tipo específico de troca epistêmica.

Este modelo foi inspirado em uma estrutura teórica conhecida como Teoria da Troca Social (TTS)¹² (Dutilh Novaes, 2020b). Trata-se de uma teoria desenvolvida por sociólogos e psicólogos sociais que busca explicar o comportamento social humano em termos de processos de troca, envolvendo custos e recompensas, e no contexto de redes sociais e estruturas de poder (Cook, 2013). Essa teoria foi originalmente desenvolvida no final da década de 1950 e início da década de 1960, sob a influência de pesquisas em economia (teoria da escolha racional), psicologia

¹¹ Não creio que a narrativa milliana seja plenamente convincente nem mesmo como uma abordagem normativa, mas uma discussão aprofundada desse ponto ultrapassa o escopo deste artigo. Para uma crítica à ideia milliana de que engajar-se com divergentes é sempre racional ou desejável, veja Fantl (2018).

¹² Nota de tradução: no original em inglês, a expressão 'Teoria da Troca Social' é abreviada como SET, da expressão língua inglesa '*Social Exchange Theory*'.

(comportamental) e trabalhos antropológicos de Malinowski, Mauss e Lévi-Strauss. A TTS é uma teoria influente e empiricamente robusta, que tem sido utilizada para investigar uma ampla gama de fenômenos sociais, como relacionamentos amorosos, interações de negócios, confiança em instituições públicas, entre muitos outros. Em particular, e de forma relevante para nossos propósitos, essa teoria tem sido amplamente utilizada para investigar a comunicação interpessoal (Roloff, 2015). Os modelos da TTS não são nem puramente descritivos — pois dependem de certas suposições idealizadas, como a de que os agentes buscam maximizar recompensas e minimizar custos — nem puramente normativos — já que incorporam descobertas experimentais, bem como dados observacionais extensivos. Além disso, a TTS combina uma perspectiva de primeira pessoa, que explica e prevê as escolhas que os indivíduos fazem entre diferentes potenciais parceiros de troca, com uma perspectiva de terceira pessoa, que foca em características estruturais dessas redes de troca.

O modelo de troca epistêmica de três níveis adapta *insights* e resultados da TTS para trocas que são especificamente *epistêmicas*, ou seja, quando recursos epistêmicos, como conhecimento, evidências, informações e assim por diante, estão envolvidos (possivelmente com outros tipos de recursos)¹³. O modelo permite uma análise meticulosa das condições sob as quais uma troca epistêmica bem-sucedida pode ocorrer ou não. Crucialmente, parece haver dois estágios preliminares que determinam se agentes específicos estarão em posição de se engajar em uma troca epistêmica profícua: as *redes* que determinam a quais fontes e recursos epistêmicos um agente é exposto; e as *escolhas contrastantes* que os agentes devem fazer sobre quais conteúdos e fontes engajar (dentre aqueles aos quais estão expostos). Visto isso, os três níveis para a troca epistêmica são:

1. **Atenção/exposição.** O primeiro nível diz respeito a saber se as pessoas são potenciais parceiras de troca umas das outras, dadas as estruturas de oportunidade relevantes para o engajamento epistêmico dentro de uma rede. Em termos mais simples: quem está na rede de contatos potenciais de um agente? Quem está em posição de atrair a atenção dos outros? Pode ser que algumas linhas de comunicação potenciais sejam cortadas, como no caso de censura estrutural ou bolhas epistêmicas. Mas também pode ser que

¹³ Veja Dutilh Novaes (2020b) para mais detalhes sobre como o modelo de três níveis emerge da TTS.

tantos sinais estejam sendo transmitidos ao ponto de várias fontes diferentes estarem competindo pela atenção do receptor (Gershberg e Illing, 2022), em uma chamada ‘economia da atenção’ (Franck, 2019).¹⁴

2. **Escolhendo com quem se engajar.** Este nível inclui as escolhas que os agentes fazem no contexto de possibilidades de troca, conforme determinadas pelas estruturas de oportunidade relevantes. Tipicamente, haverá várias opções para um determinado agente – por exemplo, os diversos jornais que posso ler em um dado dia qualquer, entre aqueles aos quais tenho acesso. Dadas limitações de tempo e atenção, escolhas contrastantes terão que ser feitas. Entre aquelas fontes que chamaram minha atenção inicial, quem vejo como digno de ser considerado um parceiro de troca? Neste ponto, considerações sobre *confiabilidade* (Hawley, 2019) e *expertise* (Goldman, 2018) entram em jogo, assim como o valor percebido do conteúdo oferecido por diferentes potenciais parceiros de troca. Em particular, confiar em alguém frequentemente implicará em *não* confiar em outra pessoa, especialmente quando suas respectivas mensagens entram em conflito (Dutilh Novaes, 2020b).
3. **Engajamento com o conteúdo.** É somente no terceiro nível que ocorre o engajamento com o *conteúdo*, propriamente dito; é aqui que a troca epistêmica real acontece. Neste ponto, o receptor se envolverá de forma reflexiva (e talvez crítica) com o argumento apresentado, buscando compreender sua substância e avaliar sua cogência. No caso de uma avaliação positiva, isso pode levar a uma mudança de visão para o receptor (embora mesmo neste ponto a receptora ainda possa hesitar em revisar suas crenças). Isso também pode levar a uma troca mutuamente benéfica, onde tanto o argumentador quanto o destinatário melhoram suas respectivas posições epistêmicas, como postulado por Mill, e em alguns casos chegam até mesmo a criar juntos novos recursos epistêmicos (como no modelo de “provas e refutações” de Lakatos sobre a prática matemática: Lakatos,

¹⁴ Veja Dutilh Novaes e de Ridder (2021) para uma discussão sobre escassez versus abundância excessiva de informações em ambientes epistêmicos.

1976).

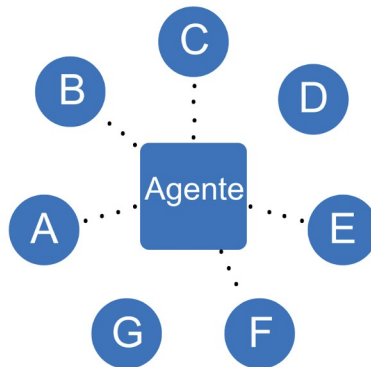


Figura 1: Atenção.

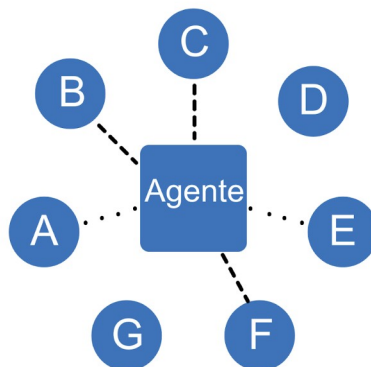


Figura 2: Escolhas.

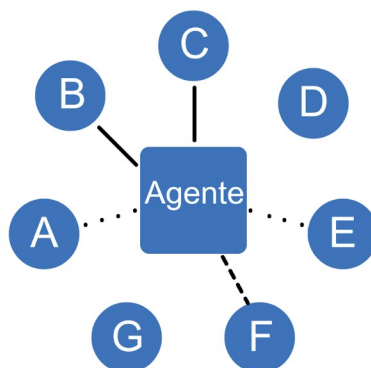


Figura 3: Engajamento.

As figuras 1 a 3 representam os três níveis.¹⁵ Para simplificar, um agente principal é representado com outros agentes ao seu redor, mas o modelo, na verdade, foca em redes complexas de agentes que estão interconectados em diferentes graus. As topologias de tais redes determinam de forma crucial como esses processos sócio-epistêmicos se desenvolvem.¹⁶

- **Atenção:** a agente não ‘vê’ as fontes D e G; as outras fontes chamam a sua atenção (linhas pontilhadas).
- **Escolhas contrastantes:** a agente considera B, C e F como dignos de troca (linhas tracejadas), mas não A e E.
- **Engajamento:** a agente eventualmente se envolve substancialmente com B e C (linhas contínuas), mas não com F.

Concepções milianas de argumentação tendem a se concentrar principalmente no nível 3 — a ‘força’ de um argumento por si só deveria ser suficiente para mudar opiniões — e a minimizar alguns dos obstáculos estruturais para uma troca verdadeiramente livre e igualitária de ideias.¹⁷ De fato, os níveis 1 e 2 determinam crucialmente se e quando alguém se envolverá seriamente com os recursos epistêmicos oferecidos por outra pessoa. Assim como os modelos originais TTS, o modelo de três níveis não é nem puramente normativo nem puramente descritivo. Não é puramente normativo porque não considera agentes ideais ou idealizados: em vez disso, considera agentes com recursos cognitivos limitados e que são suscetíveis ao que Neil Levy (2021) descreve como ‘crenças ruins’, ou seja, crenças incompatíveis com a opinião de especialistas e com nossas melhores evidências disponíveis. Além disso, o modelo é empiricamente robusto, pois se

¹⁵ O modelo também pode ser entendido em termos de conjuntos de restrições: em um determinado momento, o conjunto de pessoas com quem realmente me envolvo em trocas epistêmicas é um subconjunto daquelas que considero dignas de troca (acima de um certo limite), que, por sua vez, é um subconjunto daquelas que, devido às nossas respectivas posições na rede, são potenciais parceiros de troca para mim.

¹⁶ Observe que há uma série de semelhanças estruturais interessantes entre o modelo de três níveis que apresento aqui e o programa de pesquisa em *epistemologia de redes*, conforme desenvolvido por Zollman (2013), Olsson (2013) e O’Connor e Weatherall (2019), entre outros. Por questões de espaço, não desenvolvo este ponto mais detalhadamente aqui; ele permanecerá como um tópico para pesquisas futuras.

¹⁷ A ênfase de Mill na liberdade de expressão visa criar um ambiente informacional maximamente inclusivo e, assim, aumentar a exposição a visões distintas (fenômenos pertencentes ao nível 1). Mill menciona alguns fatores como tendo um papel central para aumentar a probabilidade de sucesso das trocas, como a importância da educação. No entanto, ele não fornece uma análise detalhada das condições sob as quais uma troca epistêmica bem-sucedida pode ocorrer ou não; em particular, trocas podem falhar mesmo em contextos em que a perseguição (religiosa ou de outra natureza) não está presente.

baseia em décadas de descobertas experimentais e observacionais da TTS pertencentes às trocas de forma mais geral. No entanto, o modelo também não é puramente descritivo ou preditivo, pois *busca explicar* os mecanismos que levam diferentes pessoas a se envolverem em trocas epistêmicas com algumas fontes, mas não com outras; isso é feito com base em alguns princípios fundamentais, como reciprocidade e justiça, destacando em particular os papéis da atenção e da confiança nesses processos. Como tal, o modelo talvez seja melhor entendido como um *modelo explanatório*, no sentido de buscar representar algumas das causas do fenômeno-alvo e os mecanismos responsáveis por provocá-lo (Ivani e Dutilh Novaes, 2022). (Ele também pode levar a recomendações *prescritivas* sobre como facilitar certos tipos de trocas epistêmicas.)

O modelo de três níveis oferece uma explicação de por que argumentos frequentemente falham em mudar opiniões, destacando algumas das condições necessárias para que isso ocorra. Primeiro, deve surgir uma adequada relação de atenção e exposição entre o emissor e o receptor — o que está longe de ser óbvio, especialmente em ambientes informacionais altamente saturados, como os que atualmente habitamos (Gershberg e Illing, 2022).

Em segundo lugar, um agente deve fazer escolhas sobre com quem se engajar, entre as diferentes possibilidades: é aqui que surgem considerações de confiabilidade — entendida como relacionada tanto à *competência* quanto à *benevolência*. (Dutilh Novaes, 2020b, 2023) Se eu já suspeito que uma determinada fonte não mantém atitudes benevolentes em relação a mim, realmente devo gastar meu precioso tempo e energia engajando-me com os seus argumentos? Talvez não (Köymen e Dutilh Novaes, no prelo). Por exemplo, a recusa de se engajar com argumentos científicos que apoiam a eficácia e segurança das vacinas por parte dos chamados ‘negacionistas das vacinas’ (‘anti-vaxxers’) é frequentemente justificada pela (não inteiramente irrazoável) suspeita de que interesses espúrios estão envolvidos (por exemplo, a narrativa da “perversa Grande Indústria Farmacêutica”: Dutilh Novaes, 2020b; Ivani e Dutilh Novaes, 2022).

Por fim, a própria troca exige que agentes com contextos epistêmicos muito diversos encontrem um terreno comum suficiente e meios adequados de comunicação, em vez de falarem sem se entender, o que está longe de ser óbvio, especialmente em situações de desacordo ideológico ou político (Talissee, 2019). Se

a (potencial) troca falhar em qualquer um desses três níveis, os argumentos não provocarão uma mudança de opinião.

EXEMPLOS DA VIDA REAL

Apesar de todos esses desafios, a conclusão de que argumentos *nunca* mudam opiniões também é infundada: de fato, às vezes, argumentos *realmente* mudam opiniões. Diante disso, a questão que se coloca é a seguinte: em quais condições isso é (mais) provável de acontecer? O modelo de três níveis fornece ferramentas conceituais adequadas para abordar essa questão. Em vez de discuti-la de forma abstrata, ou com exemplos artificiais, apresento dois exemplos concretos recentes de pessoas que passaram por transformações epistêmicas radicais em que argumentos (presumivelmente) desempenharam um papel significativo: Megan Phelps-Roper, um ex-membro de destaque da Igreja Batista de Westboro,¹⁸ e Derek Black, um ex-defensor destacado da supremacia branca.¹⁰

A Igreja Batista de Westboro é uma congregação hiper-calvinista com sede em Topeka, Kansas, frequentemente descrita como um grupo de ódio. Tal igreja é conhecida por realizar inflamadas manifestações homofóbicas e por produzir discursos de ódio contra ateus, judeus, muçulmanos, pessoas trans e inúmeras denominações cristãs. Megan Phelps-Roper é neta do fundador Fred Phelps e foi criada para ser um membro proeminente do grupo. Nesse ambiente, ela cresceu imersa em sua rígida ideologia e, desde cedo, participou de manifestações em funerais de homens gays (com cartazes contendo frases como ‘DEUS ODEIA VIAD*S’) e, posteriormente, de soldados mortos em guerra (já que os membros da Westboro acreditam que as guerras em que os EUA se envolveram nas últimas décadas são resultados de um castigo de Deus pela tolerância do país à homossexualidade).

Apesar das posições extremas dos membros de Westboro, seus filhos, incluindo Megan, em geral, frequentavam as escolas públicas de Topeka. Na escola, Megan foi presumivelmente exposta a outras visões de mundo mais tolerantes, mas isso não afetou substancialmente sua própria convicção no sistema de crenças de

¹⁸ Minha discussão sobre a trajetória de Megan Phelps-Roper baseia-se principalmente no perfil dela publicado pela *New Yorker* em 2015 (Chen, 2015). Ela também escreveu uma autobiografia, apropriadamente intitulada ‘*Unfollow: A Journey from Hatred to Hope*’ (Phelps-Roper, 2020).

¹⁰ Minha discussão sobre a trajetória de Derek Black baseia-se principalmente em uma entrevista para o podcast ‘*The Daily*’ do *New York Times* (Stone, 2017). Há também um livro narrando a jornada de Black, escrito pelo jornalista Eli Saslow (2018).

Westboro. Muitos membros tinham ensino superior, e alguns, incluindo a mãe de Megan, trabalhavam como advogados. Assim, eles não viviam exatamente em uma bolha epistêmica no sentido de não serem expostos a sistemas de crenças alternativos; de fato, acreditavam que os membros de Westboro poderiam pregar melhor para os "ímpios" vivendo entre eles. A ideia era: se você realmente conhece a "verdade" em seu coração, a exposição ao mundo dos ímpios não afetará sua devoção. Na prática, no entanto, não havia espaço algum para autonomia epistêmica ou divergência: o valor supremo incutido nas crianças era o da total obediência.

Em 2009, Megan entrou no Twitter para espalhar ainda mais as ideias de Westboro. Alguns de seus *tweets* homofóbicos foram notados e retuitados (no tom de ridicularização) por contas grandes, o que resultou em muitas respostas indignadas, mas também em um número significativo de seguidores. Com isso, ela passou a entrar em contato com uma gama mais ampla de críticas, aos quais respondia diligentemente citando passagens bíblicas (junto com referências à cultura pop e emojis). Ela estava acostumada a dar entrevistas a jornalistas, mas no Twitter podia interagir diretamente com muitas pessoas, sem um filtro jornalístico.

Porém, naquela época, Megan também havia começado a ter dúvidas sobre alguns dos ensinamentos de Westboro. Em particular, por volta do momento em que entrou no Twitter, Westboro estava se preparando para o fim do mundo. Havia previsões muito específicas sobre como os membros de Westboro conduziriam cento e quarenta e quatro mil judeus que se arrependessem por terem matado Jesus pelo deserto de Israel, até que Cristo finalmente viesse salvá-los. Megan sentia que não havia suporte adequado nas escrituras para muitas dessas previsões e recorreu ao Twitter em busca de respostas. Mais especificamente, ela começou a seguir e se envolver com alguns usuários judeus do Twitter, em particular com um web designer de Jerusalém chamado David Abitbol.

E assim, mesmo que por meio de debates acirrados, ela começou a estabelecer conexões mais profundas com outros usuários do Twitter. Até então, as interações com "os ímpios" haviam permanecido superficiais e passageiras, como as que ela tinha com os opositores em manifestações. No entanto, no Twitter, ela se envolveu em debates prolongados com pessoas específicas (como Abitbol), com quem

desenvolveu padrões de interação acalorados, mas amigáveis. Para sua surpresa, pela primeira vez em sua vida, ela começou a *se importar* com o que pessoas de fora de Westboro pensavam dela — particularmente, alguns de seus conhecidos do Twitter; as conexões com alguns de seus interlocutores do Twitter se tornaram cada vez mais significativas. (De fato, ela acabou se casando com um deles anos depois.)

E assim, como resultado de pequenas sementes de dúvida em relação à pregação de Westboro (bem como preocupações relacionadas a mudanças na forma como a igreja era administrada e ao papel das mulheres nela), mas principalmente por meio de suas conexões e interações no Twitter, Megan embarcou em um longo e doloroso processo de questionar tudo o que havia sido ensinada a acreditar. Cerca de três anos após entrar no Twitter, ela começou a considerar seriamente deixar a igreja. Isso, é claro, implicaria em enormes custos sociais e emocionais; ela basicamente perderia todo o contato com sua família imediata e estendida. Consequentemente, ela tomou a decisão de deixar a igreja (junto com sua irmã mais nova, Grace) em novembro de 2012 e começou a se reconectar, inclusive em ambientes offline, com alguns de seus contatos no Twitter, como Abitbol. Desde então, ela se tornou uma defensora do diálogo entre grupos com visões conflitantes e tem discursado em diversos locais sobre suas experiências (incluindo o inevitável *TED Talk*).

Se acreditarmos em seu próprio relato sobre o processo, os argumentos desempenharam um papel importante no “avanço epistêmico” (lento, mas profundo) de Megan ao perceber que não podia mais endossar o sistema de crenças de Westboro. Houve muita deliberação envolvida, tanto consigo mesma quanto com muitos de seus contatos no Twitter (alguns dos quais também eram conhecedores de fontes que ela considerava confiáveis, em particular a Bíblia). Por meio desses processos (que, em alguns momentos, se assemelhavam a diálogos socráticos), as inconsistências e as discrepâncias nas doutrinas de Westboro tornaram-se evidentes para ela, levando a uma revisão completa de suas próprias convicções.

No entanto, duas condições necessárias precisavam estar presentes para que esses argumentos funcionassem: naturalmente, ela precisava ser exposta a eles (no Twitter, ela podia ser exposta a uma ampla gama de fontes e interlocutores); porém, mais importante, esses argumentos estavam vindo de pessoas a quem ela havia passado a respeitar e se importar. Ela já tinha sido exposta a ideias que conflitavam com as doutrinas de Westboro antes (por exemplo, na escola), mas, desta

vez, as fontes dessas ideias eram pessoas com quem ela havia estabelecido conexões mais profundas. Desta vez, ela prestou mais atenção e se envolveu seriamente com o conteúdo dos argumentos. O que talvez seja notável na trajetória de Megan é o fato de que o processo de recalibração das atribuições de respeito e confiança a diferentes pessoas (afastando-se dos membros de Westboro e voltando-se para os “de fora”) ocorreu principalmente por meio de interações online, em vez de presenciais. (Seu perfil na revista *New Yorker* descreve o processo como “conversão via Twitter”.) Conexões online podem, afinal, se tornar conexões “reais” e podem oferecer uma rede muito mais ampla de potenciais parceiros de troca epistêmica.¹¹

A trajetória de Derek Black apresenta semelhanças interessantes com a de Megan Phelps-Roper, mas, no caso dele, a “conversão” ocorreu principalmente por meio de interações presenciais, e não em um contexto online. Derek é filho de Don Black, proeminente supremacista branco e fundador do *Stormfront*, uma das comunidades online de supremacistas brancos mais influentes dos EUA. Seu padrinho é David Duke, um dos líderes mais visíveis e influentes da Ku Klux Klan nas últimas décadas (como mostrado, por exemplo, no filme de Spike Lee de 2018, *Infiltrado na Klan*).¹² Tanto Duke quanto Don Black são Grandes Magos da Ku Klux Klan. Derek foi criado para ser o “príncipe herdeiro” da supremacia branca nos Estados Unidos e, desde cedo, esteve profundamente envolvido na promoção dessa visão de mundo, incluindo a produção de um programa de rádio com seu pai.

Diferentemente de Megan Phelps-Roper, Derek foi educado no modelo de ensino doméstico (*homeschooling*) e, portanto, durante a sua juventude teve exposição limitada a visões de mundo que não fossem as crenças supremacistas brancas de sua família. Todo o seu mundo socioemocional enquanto crescia era composto por pessoas que defendiam a mesma ideologia. Em 2010, aos 21 anos, ele decidiu se matricular no *New College of Florida*, em Sarasota, a quatro horas de carro de sua casa; essa foi a primeira vez que ele deixou o mundo insular da supremacia branca em que havia crescido. Ele começou a viver o que poderia ser descrito como uma “vida dupla”: gravando um programa de rádio para promover ideias supremacistas brancas com seu pai pela manhã e, durante o resto do dia, frequentando aulas e socializando com estudantes que eram defensores (de esquerda) da justiça so-

¹¹ Veja Lewiński e Dutilh Novaes (no prelo) para uma análise da comunicação online baseada no modelo de três níveis de troca epistêmica.

¹² O nome original do filme, em inglês, é *BlackKkKlansman*.

cial. Inicialmente, sua identidade como supremacista branco não havia sido revelada.

Todavia, inevitavelmente, após alguns meses, um colega descobriu e expôs sua identidade, e suas crenças racistas e ativismo contínuo tornaram-se de conhecimento público na faculdade. Não surpreendentemente, isso levou ao seu ostracismo entre os estudantes. A única exceção foi um pequeno grupo de estudantes judeus que começou a convidá-lo para seus jantares de *Shabbat*. (Nessa época, Derek já havia tido um breve relacionamento com uma mulher judia, que chegou ao fim quando sua persona supremacista branca se tornou de conhecimento público.) Talvez porque essas fossem as únicas pessoas ainda dispostas a socializar com ele, ele se tornou um frequentador regular de seus jantares.

Enquanto alguns dos participantes dos jantares não buscavam confrontar Derek abertamente sobre suas crenças, outros se envolviam em acaloradas discussões intelectuais com ele. Aqui está o próprio relato dele sobre essas discussões:

Eu dizia: “É isso que eu acredito sobre diferenças de QI, tenho 12 estudos diferentes que foram publicados ao longo dos anos, aqui está a revista que reuniu essas informações, eu acredito que isso é verdade, que a raça prediz o QI e que existem diferenças de QI entre as raças.” E eles retornavam com 150 estudos mais recentes e mais bem pesquisados e me explicavam como as estatísticas funcionam, e íamos e voltávamos até que eu chegasse ao fim daquele argumento e dissesse: Sim, isso faz sentido, isso não se sustenta e eu vou remover isso da minha caixa de ferramentas ideológicas, mas todo o resto ainda está lá. E fizemos isso ao longo de um ou dois anos, uma coisa após outra, até que cheguei a um ponto em que não acreditava mais nisso. (Stone, 2017)

Essas conversas duraram anos, durante os quais Derek gradualmente se afastou da ideologia supremacista branca com a qual havia crescido. Eventualmente, em 2013, Derek escreveu uma declaração pública para o Southern Poverty Law Center, renunciando publicamente às suas visões anteriores. Ele tinha muito a perder, do ponto de vista social e emocional, ao se distanciar da

supremacia branca, incluindo seu relacionamento próximo com sua família: mudar de opinião pode ser não apenas cognitivamente, mas também socialmente custoso (um aspecto também explorado em Gordon-Smith, 2019). Assim como no caso de Megan Phelps-Roper, o avanço epistêmico de Derek não aconteceu da noite para o dia: foi o resultado de um longo processo em que suas crenças originais foram gradualmente dissipadas, ao menos parcialmente pela força dos argumentos (isto é, se acreditarmos no próprio relato dele sobre esse processo).

No entanto, mais uma vez, o fato de os argumentos virem de pessoas que Derek havia passado a respeitar em nível pessoal (apesar de as etnias de algumas delas serem consideradas “inferiores” segundo a visão supremacista branca que ele defendia até então) foi um elemento crucial no processo. Ele realmente ouviu e se envolveu com o conteúdo dos argumentos por conta desse ambiente interpessoal favorável, o que, por sua vez, foi facilitado por sua vulnerabilidade e pelo fato de essas serem as únicas pessoas ainda dispostas a interagir com ele no campus. (Assim como Megan, Derek também acabou em um relacionamento amoroso de longo prazo com uma das pessoas que desafiaram suas crenças no início.) Até ir para o *New College*, a exposição de Derek a outras visões de mundo havia sido limitada (um fenômeno de nível 1), e ele havia sido criado para confiar apenas naqueles que defendiam ideias semelhantes às de sua família (um fenômeno de nível 2). Nos termos de Nguyen (2020), ele estava tanto em uma bolha epistêmica quanto em uma câmara de eco, enquanto Megan Phelps-Roper estava principalmente em uma câmara de eco, mas não tanto em uma bolha epistêmica.¹³ A reconfiguração dos circuitos de atenção e confiança desencadeada por suas experiências no campus foram os elementos que permitiram que os argumentos mudassem a opinião de Derek.

Naturalmente, argumentos também podem mudar opiniões em questões mais específicas. Os dois casos descritos aqui correspondem a revisões completas de sistemas de crenças inteiros, mas argumentos também podem, e provavelmente com mais facilidade, causar revisões localizadas (que podem exigir algumas adaptações, mas não de forma tão radical quanto nesses dois casos). O ponto aqui é que, se *mesmo* nesses dois casos extremos os argumentos parecem ter mudado as

¹³ ‘Uma bolha epistêmica é uma estrutura social epistêmica na qual outras vozes relevantes foram deixadas de fora, talvez acidentalmente. Uma câmara de eco é uma estrutura social epistêmica da qual outras vozes relevantes foram ativamente excluídas e desacreditadas’ (Nguyen, 2020, p. 141).

opiniões de Megan e Derek, então, a *fortiori*, argumentos podem também provocar revisões de crenças mais localizadas.

Um tema que investiguei em trabalhos anteriores é a mudança na opinião pública em relação ao personagem folclórico ‘Black Pete’ nos Países Baixos (Zwarte Piet) (Dutilh Novaes et al., 2020). Black Pete é o assistente de São Nicolau nas festividades extremamente populares de São Nicolau no início de dezembro, e era tradicionalmente retratado de maneiras altamente racializadas (blackface, cabelo encaracolado, lábios grossos e vermelhos). Nos últimos anos, houve uma mudança significativa na opinião pública sobre a suposta natureza racista do personagem; enquanto há cerca de dez anos o personagem era visto positivamente por 95% da população, atualmente pelo menos um terço da população (e crescendo) passou a considerá-lo inaceitável. Isso levou a mudanças importantes na forma como o personagem é retratado; com destaque na queda acentuada no uso de maquiagem de blackface. Os argumentos parecem ter desempenhado um papel importante nessa mudança de opinião pública, particularmente ao desafiar o que Charles Mills descreveu de forma precisa como “ignorância branca”: uma perspectiva epistêmica deformante que torna pessoas classificadas como brancas alheias à abrangência do racismo (Mills, 2015).

ESCLARECIMENTOS

Antes de concluir, são necessários alguns esclarecimentos sobre o quadro delineado até agora. Em primeiro lugar, de uma perspectiva progressista-igualitária, as “conversões” de Megan e Derek são vistas como positivas, porque eles renunciaram ao que muitos de *nós* consideramos como visões de mundo erradas e problemáticas. Eles atingiram o que consideramos estados epistêmicos significativamente melhores.¹⁴ Mas os mecanismos gerais descritos pelo modelo de três níveis — caminhos de atenção, confiança e engajamento — não favorecem ideologias específicas (Dutilh Novaes, 2023). De fato, a disseminação de posições “problemáticas”, como a rejeição às vacinas e várias teorias da conspiração, segue padrões semelhantes. Em particular, a propensão a adotar teorias da conspiração parece estar fortemente associada à desconfiança em relação a instituições

¹⁴ Pelo menos, estou assumindo que a maioria dos leitores deste texto rejeitará a homofobia, o racismo e a supremacia branca.

estabelecidas, como governos, a grande imprensa e a comunidade científica (van Prooijen, Spadaro e Wang, 2022). No final das contas, se passamos a adotar crenças “boas” ou “ruins” (na terminologia de Levy) é, em grande medida, resultado dos ambientes epistêmicos em que nos encontramos, incluindo nossas atribuições de credibilidade e confiabilidade a diferentes fontes (Dutilh Novaes, 2023). Os processos de formação de crenças que levam a crenças “boas” ou “ruins” não são, talvez de forma surpreendente, fundamentalmente diferentes (Levy, 2021). Crenças ruins também podem ser sustentadas por argumentos — argumentos “ruins”, talvez (embora não necessariamente!), mas ainda assim argumentos — e eles também podem mudar opiniões se as condições forem adequadas. (Observe que este também é um ponto delicado para o otimista milliano que acredita que a verdade eventualmente prevalecerá, desde que todas as opiniões possam ser expressas e discutidas abertamente.)

Um segundo esclarecimento diz respeito a se existem condições conjuntamente *suficientes* para que argumentos mudem opiniões. O que o modelo de três níveis descreve são condições *necessárias* relacionadas à atenção e às atribuições de credibilidade. Mas, mesmo que essas condições estejam presentes, não há garantia de que os argumentos realmente mudarão opiniões; as pessoas podem, e de fato é o que frequentemente acontece, continuar apegadas a suas crenças prévias, especialmente aquelas profundamente entrelaçadas com seus modos de vida.¹⁵ Embora a ideia de um método infalível para mudar opiniões por meio de argumentos de alta qualidade possa parecer atraente, na prática, os argumentos por si só não podem *forçar* que uma atualização epistêmica aconteça, por mais que tentem.¹⁶

¹⁵ Compare a noção de *technē cultural* de Sally Haslanger, entendida como um conjunto de significados sociais ‘que fornece um “cenário” para a ação e é parte constitutiva do sistema de regulação social local’ (Haslanger, 2021, p. 23). Uma *technē cultural* ‘que deu errado’ organizará estruturas sociais de maneiras injustas e, para Haslanger, isso é exatamente o que *ideologia* é. Para desmantelar uma *technē cultural* ‘que deu errado’, argumentos racionais por si só terão pouco ou nenhum efeito; em vez disso, a *technē cultural* em questão deve primeiro ser ‘desestabilizada’ para abrir possibilidades de contestação. Nos casos de Megan e Derek, a desestabilização em questão foi causada por estarem em diferentes ambientes discursivos e afetivos (Twitter para Megan, faculdade para Derek). Mas, em um nível mais amplo, societal, desestabilizações mais significativas parecem necessárias, seguindo a noção de *technē cultural* de Haslanger. No entanto, elas ainda podem envolver parcialmente argumentos, por exemplo, no que é conhecido na tradição marxista como *conscientização*.

¹⁶ Este é um ponto relacionado ao que alguns autores identificam como a natureza intrinsecamente *coercitiva* dos argumentos (Nozick, 1981; Casey, 2020), que, no entanto, nem sempre é eficaz. Como Wittgenstein apontou, mesmo uma prova matemática correta pode falhar em persuadir, apesar da “dureza do dever lógico” (ver Wright, 1990). Uma questão relacionada, não abordada aqui, é se é eticamente aceitável tentar mudar a opinião de alguém (seja por meio de argumentos ou outras

Nesse sentido, mesmo quando ocorre uma mudança de opinião aparentemente provocada por argumentos, pode muito bem ser que as causas eficazes estejam, em última análise, relacionadas a fatores não epistêmicos, como dinâmicas sociais (por exemplo, o desejo de pertencer a um determinado grupo) ou incentivos econômicos. Em outras palavras, não podemos ter certeza de que os argumentos foram persuasivos pelos motivos “certos” (racionais), ou seja, aqueles relacionados à sua qualidade enquanto argumentos.¹⁷ De fato, dada a propensão humana à *racionalização* (Cushman, 2020), muitas vezes não é transparente para o próprio agente o que exatamente motivou uma mudança de opinião.

Por fim, Megan e Derek foram (presumivelmente) influenciados por argumentos apenas porque previamente aceitavam as regras básicas do jogo de linguagem da argumentação.¹⁸ Megan era habilidosa na prática de argumentar em apoio a crenças religiosas, com base em uma análise cuidadosa das escrituras, e passou a respeitar o conhecimento bíblico de alguns de seus interlocutores no Twitter. Derek se referia a estudos “científicos” para sustentar suas visões (por exemplo, de que existem diferenças de QI entre raças), mas depois percebeu que havia estudos científicos muito *melhores* apoiando visões opostas. Eles, portanto, eram receptivos à prática de apoiar posições com argumentos e evidências; isto é, havia pelo menos um certo grau de concordância em um meta-nível sobre as “regras do jogo” estabelecidas entre eles e seus interlocutores. Caso isso não fosse verdade — por exemplo, se eles pensassem que tudo se resumia a “o que parece certo para mim” ou que todas as opiniões são igualmente válidas — os argumentos talvez não tivessem exercido tanto impacto sobre eles.

Em suma, os casos recém-discutidos ainda não oferecem total garantia ao otimista milliano de que, sob as condições adequadas, bons argumentos de fato mudarão opiniões para melhor. Argumentos “ruins” também podem mudar opiniões para pior; bons argumentos podem falhar em provocar uma mudança de opinião, mesmo sob as circunstâncias corretas; mesmo quando parece que uma mudança de opinião foi causada pelo engajamento com argumentos de alta qualidade, não podemos ter certeza de que as causas reais para a mudança foram verdadeiramente argumentativas; e argumentos por si só terão pouco ou nenhum efeito em casos em que as pessoas rejeitam a própria ideia de atualizar suas intervenções); isso não constituiria uma violação problemática da autonomia intelectual de alguém?

¹⁷ Agradeço a James Owen Weatherall por levantar esses dois pontos.

¹⁸ Devo esse ponto a Harvey Siegel.

crenças em vista de argumentos e evidências. Ainda assim, os casos discutidos oferecem pelo menos uma explicação plausível dos mecanismos pelos quais bons argumentos podem mudar opiniões e, assim, corroboram parcialmente a visão de Mill de que envolver-se com divergentes pode permitir a correção de erros.

CONCLUSÕES

Começamos com a visão de que argumentos podem mudar opiniões apenas pela força da razão. Na prática, no entanto, e certamente em situações em que valores e visões políticas desempenham um papel significativo, argumentos não parecem ser particularmente adequados para mudar opiniões; ao contrário, as pessoas tipicamente ou definitivamente se recusam a se envolver com argumentos que conflitam com suas crenças profundamente enraizadas, ou então não são influenciadas por eles. Porém, em algumas circunstâncias, argumentos podem de fato ter sucesso em mudar opiniões. Defendi que dois fatores importantes, mas frequentemente subestimados, atenção e confiança, precisam ser levados em conta para explicar o caráter persuasivo (ou falta dele) de argumentos em situações específicas. Argumentos só podem mudar opiniões se captarem a atenção do receptor e se o receptor escolher considerá-los cuidadosamente, o que, por sua vez, é significativamente (mas não completamente) determinado por atribuições de credibilidade e confiabilidade à fonte. Se essas condições estiverem presentes, pode muito bem acontecer (embora, novamente, sem garantia!) que os argumentos mudem a opinião de alguém.¹⁹

Examinamos os casos reais de Megan Phelps-Roper e Derek Black. Em ambos os casos, eles renunciaram às visões de mundo com as quais foram criados, em partes graças ao engajamento argumentativo (ao menos, se acreditarmos em seus próprios relatos desses processos). No entanto, ambos primeiramente passaram a respeitar as *fontes* desses argumentos, e foi por isso que se envolveram seriamente com o conteúdo deles, em vez de rejeitá-los imediatamente. Além disso, em ambos os casos, foi um processo lento: argumentos precisam de tempo para realmente

¹⁹ McIntyre (2021, p. 73) apresenta uma visão muito semelhante de como negacionistas da ciência (às vezes) mudam de opinião diante de argumentos e evidências: ‘Todas essas histórias são basicamente as mesmas. Elas acontecem no contexto de um relacionamento pessoal e de confiança. Como venho dizendo desde o início, fatos e evidências podem importar, mas precisam ser apresentados pela pessoa certa no contexto certo.’

mudar opiniões. Foi o efeito cumulativo de muitas dessas interações argumentativas que eventualmente levou a um completo abandono de suas posições originais. Quão representativos esses dois casos são em uma escala mais ampla é difícil de determinar; mas, como a principal afirmação deste artigo é meramente existencial – argumentos podem, às vezes, mudar opiniões – eles oferecem suporte a essa modesta afirmação e ajudam a ilustrar os mecanismos envolvidos.

Esses dois casos também mostram que mudar opiniões por meio de argumentos é custoso (Casey, 2020). Pode ser custoso para a pessoa que muda de opinião, pois isso pode implicar na perda de suas conexões sociais e afetivas mais significativas; e é custoso para aqueles que tentam mudar opiniões por meio de argumentos, já que precisam investir recursos significativos (tempo, energia) para captar a atenção do receptor e ganhar confiança suficiente para que ele se engaje seriamente com o conteúdo dos argumentos. Além disso, o próprio processo argumentativo pode ser lento e exigir muitas iterações. Assim, uma conclusão plausível a ser tirada é que argumentos não são ferramentas muito eficientes para mudar opiniões (ao contrário, talvez, de narrativas ou discursos propagandistas). Ainda assim, não precisamos ir tão longe a ponto de concluir que argumentos são inúteis e fúteis, como alguns gostam de dizer; pelo menos nas circunstâncias certas, eles podem de fato mudar opiniões pelas razões “certas”.²⁰

REFERÊNCIAS

- BELL, D. E.; RAIFFA, H.; TVERSKY, A.. Descriptive, normative, and prescriptive interactions in decision making. In: BELL, D. E.; RAIFFA, H.; TVERSKY, A. (eds.). **Decision making: descriptive, normative, and prescriptive interactions**. Cambridge: Cambridge University Press, 1988. p. 9-30.
- CASEY, J. Adversariality and argumentation. **Informal Logic**, v. 40, n. 1, p. 77-108, 2020. Disponível em: https://informallogic.ca/index.php/informal_logic/article/view/5969/5076. Acesso em: 1 mar. 2025.

²⁰ Esta pesquisa foi generosamente apoiada pelo Conselho Europeu de Pesquisa com a concessão ERC-2017-CoG 771074 para o projeto ‘The Social Epistemology of Argumentation’. Agradecimentos à equipe SEA por suas inestimáveis contribuições ao longo dos anos. Sou grata pelo feedback recebido de várias audiências: o workshop ‘Epistemic Breakthroughs’ em St Andrews, a Conferência Europeia sobre Argumentação em Roma, a Conferência da Sociedade Brasileira de Filosofia Analítica no Rio de Janeiro e os seminários na Ruhr University Bochum e na VU Amsterdam.

CHEN, A. Conversion via Twitter. **The New Yorker**, Nova York, 15 nov. 2015. Disponível em: <https://www.newyorker.com/magazine/2015/11/23/conversion-via-twitter-westboro-baptist-church-megan-phelps-roper>. Acesso em: 1 mar. 2025.

COHEN, D. H. Argument is war... and war is hell: philosophy, education, and metaphors for argumentation. **Informal Logic**, v. 17, n. 2, p. 177-188, 1995. Disponível em: https://informallogic.ca/index.php/informal_logic/article/view/2406/1848. Acesso em: 1 mar. 2025.

COOK, K. S.; CHESHIRE, C.; RICE, E. R. W.; NAKAGAWA, S. Social exchange theory. In: DELAMATER, J.; WARD, A. (eds.). **Handbook of social psychology**. Dordrecht: Springer, 2013. p. 61-88. DOI: https://doi.org/10.1007/978-94-007-6772-0_3.

COPPOCK, A. **Persuasion in parallel: how information changes minds about politics**. Chicago; London: University of Chicago Press, 2022.

CUSHMAN, F. Rationalization is rational. **Behavioral and Brain Sciences**, v. 43, p. 1-69, 2020. Disponível em: <https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/abs/rationalization-is-rational/2A13B99ED09BD802C0924D3681FEC55B>. Acesso em: 1 mar. 2025.

DUTILH NOVAES, C. **The dialogical roots of deduction: historical, cognitive, and philosophical perspectives on reasoning**. Cambridge: Cambridge University Press, 2020a.

DUTILH NOVAES, C. The role of trust in argumentation. **Informal Logic**, v. 40, n. 2, p. 205-236, 2020b. Disponível em: https://informallogic.ca/index.php/informal_logic/article/view/6328. Acesso em: 1 mar. 2025.

DUTILH NOVAES, C. **Who's afraid of adversariality? Conflict and cooperation in argumentation**. *Topoi*, v. 40, p. 873-886, 2021. Disponível em: <https://rdcu.be/ebK5U>. Acesso em: 1 mar. 2025. DOI: <https://doi.org/10.1007/s11245-020-09736-9>.

DUTILH NOVAES, C. The (higher-order) evidential significance of attention and trust –comments on Levy's Bad Beliefs. **Philosophical Psychology**, 2023. Disponível em: <https://doi.org/10.1080/09515089.2023.2174845>. Acesso em: 1 mar. 2025.

DUTILH NOVAES, C.; DE RIDDER, J. Is fake news old news? In: BERNECKER, S.; FLOWERREE, A. K.; GRUNDMANN, T. (eds.). **The epistemology of fake news**. Oxford: Oxford University Press, 2021. Disponível em: <https://doi.org/10.1093/oso/9780198863977.003.0008>. Acesso em: 28 fev. 2025.

DUTILH NOVAES, C.; SULLIVAN, E.; LAGEWAARD, T.; ALFANO, M. Changing minds through argumentation: Black Pete as a case study. In: DUTILH NOVAES, C.; JANSEN, H.; VAN LAAR, J. A.; VERHEIJ, B. (eds.). **Reason to dissent: proceedings of the 3rd European Conference on Argumentation, Volume II**. Londres: College Publications, 2022.

2020. v. 2, p. 243-260. Disponível em: https://research.vu.nl/files/256191602/Changing_minds_through_argumentation.pdf. Acesso em: 1 mar. 2025.

FANTL, J. **The limitations of the open mind**. Oxford: Oxford University Press, 2018.

FOGELIN, R. J. The logic of deep disagreements. **Informal Logic**, v. 7, n. 1, p. 1-8, 1985. Reimpresso em **Informal Logic**, v. 25, n. 1, p. 3-11, 2005. Disponível em: https://informallogic.ca/index.php/informal_logic/article/view/1040/635. Acesso em: 1 mar. 2025.

FRANCK, G. The economy of attention. **Journal of Sociology**, v. 55, n. 1, p. 8-19, 2019. Disponível em: <https://doi.org/10.1177/1440783318811778>. Acesso em: 1 mar. 2025.

GERSHBERG, Z.; ILLING, S. D. **The paradox of democracy: free speech, open media, and perilous persuasion**. Chicago: University of Chicago Press, 2022.

GOLDMAN, A. I. Argumentation and social epistemology. **The Journal of Philosophy**, v. 91, n. 1, p. 27-49, jan. 1994. Disponível em: <https://doi.org/10.2307/2940949>. Acesso em: 1 mar. 2025.

GOLDMAN, A. I. An epistemological approach to argumentation. **Informal Logic**, v. 23, n. 1, p. 51-63, 2003. Disponível em: https://informallogic.ca/index.php/informal_logic/article/view/2153. Acesso em: 1 mar. 2025.

GOLDMAN, A. I. Expertise. **Topoi**, v. 37, p. 3-10, 2018. Disponível em: <https://doi.org/10.1007/s11245-016-9410-3>. Acesso em: 1 mar. 2025.

GORDON-SMITH, E. **Stop being reasonable: how we really change minds**. New York: Public Affairs, 2019.

GRIMM, S. R.; BAUMBERGER, C.; AMMON, S. **Explaining understanding: new perspectives from epistemology and philosophy of science**. London: Routledge, 2016.

HALLIDAY, D.; McCABE, H. John Stuart Mill on free speech. In: COADY, D.; CHASE, J. (eds.). **The Routledge handbook of applied epistemology**. London: Routledge, 2019. p. 72-87.

HANNON, M. Empathetic understanding and deliberative democracy. **Philosophy and Phenomenological Research**, v. 101, n. 3, p. 591-611, 2019. Disponível em: <https://doi.org/10.1111/phpr.12624>. Acesso em: 1 mar. 2025.

HASLANGER, S. Political epistemology and social critique. **Oxford Studies in Political Philosophy**, v. 7, 2021. Disponível em: <https://doi.org/10.1093/oso/9780192897480.003.0002>. Acesso em: 1 mar. 2025.

HAWLEY, Katherine. **How to be trustworthy**. Oxford: Oxford University Press, 2019.

ISENBERG, D. J. Group polarization: a critical review and meta-analysis. **Journal of Personality and Social Psychology**, v. 50, n. 6, p. 1141-1151, 1986. Disponível em: <https://doi.org/10.1037/0022-3514.50.6.1141>. Acesso em: 1 mar. 2025.

IVANI, S.; DUTILH NOVAES, C. Public engagement and argumentation in science. **European Journal for Philosophy of Science**, v. 12, n. 54, 2022. Disponível em: <https://doi.org/10.1007/s13194-022-00480-y>. Acesso em: 1 mar. 2025.

KAHAN, D. M. Misconceptions, misinformation, and the logic of identity-protective cognition. **Social Science Research Network**, 2017. Disponível em: <https://ssrn.-com/abstract=2973067>. Acesso em: 1 mar. 2025.

KÖYMEN, B.; DUTILH NOVAES, C. Reasoning and trust: a developmental perspective. In: STOVALL, Preston; KOREŇ, Ladislav (eds.). **Why and how we give and ask for reasons**. Oxford: Oxford University Press, [no prelo].

KUHN, T. S. **The structure of scientific revolutions**. 4. ed. Chicago; London: University of Chicago Press, 2012.

LAKATOS, I. **Proofs and refutations: the logic of mathematical discovery**. Cambridge: Cambridge University Press, 1976.

LAUGHLIN, P. R. **Group problem solving**. Princeton: Princeton University Press, 2011.

LEVY, N. **Bad beliefs: why they happen to good people**. New York: Oxford University Press, 2021.

LEWIŃSKI, M.; DUTILH NOVAES, C. The many-to-many model: communication, attention, and trust in online conversations. In: CONNOLLY, P.; GOLDBERG, S. C.; SAUL, J. (eds.). **Conversations online**. Oxford: Oxford University Press, [no prelo].

LONGINO, H. **Science as social knowledge: values and objectivity in scientific inquiry**. Princeton: Princeton University Press, 1990.

MCINTYRE, L. C. **How to talk to a science denier: conversations with flat earthers, climate deniers, and others who defy reason**. Cambridge: MIT Press, 2021.

MERCIER, H. Reasoning and argumentation. In: BALL, L. J.; THOMPSON, V. A. (eds.). **The Routledge international handbook of thinking and reasoning**. London: Routledge/Taylor & Francis Group, 2018. p. 401-414.

MERCIER, H. **Not born yesterday: the science of who we trust and what we believe**. Princeton: Princeton University Press, 2020.

MILL, J. S. **On liberty**. [1859]. Peterborough: Broadview Press, 1999.

MILL, J. S. **Sobre a liberdade**. Tradução Pedro Madeira. Rio de Janeiro: Nova Fronteira, 2011.

MILLS, C. W. Global white ignorance. In: GROSS, Matthias; MCGOEY, L. (eds.). **Routledge international handbook of ignorance studies**. London: Routledge, 2015. p. 217-227.

NGUYEN, C. T. Echo chambers and epistemic bubbles. **Episteme**, v. 17, n. 2, p. 141-161, 2020. Disponível em: <https://doi.org/10.1017/epi.2018.32>. Acesso em: 1 mar. 2025.

NICKERSON, R. S. Confirmation bias: a ubiquitous phenomenon in many guises. **Review of General Psychology**, v. 2, n. 2, p. 175-220, 1998. Disponível em: <https://doi.org/10.1037/1089-2680.2.2.175>. Acesso em: 1 mar. 2025.

NOZICK, R. **Philosophical explanations**. Cambridge: Harvard University Press, 1981.

O'CONNOR, C.; WEATHERALL, J. O. **The misinformation age: how false beliefs spread**. New Haven: Yale University Press, 2019.

OLSSON, E. J. A Bayesian simulation model of group deliberation and polarization. In: ZENKER, F. (ed.). **Bayesian argumentation**. Dordrecht: Springer, 2013. v. 362. p. 113-133. Disponível em: https://doi.org/10.1007/978-94-007-5357-0_6. Acesso em: 1 mar. 2025.

PHELPS-ROPER, M. **Unfollow: a journey from hatred to hope**. London: Riverrun, 2020.

QUINE, W. V. O. Two dogmas of empiricism. **Philosophical Review**, v. 60, n. 1, p. 20-43, 1951.

ROLOFF, M. E. Social exchange theories. In: BERGER, C. R.; ROLOFF, M. E. (eds.). **The international encyclopedia of interpersonal communication**. v. 3. London: Wiley, 2015. p. 1615-1633.

SASLOW, E. **Rising out of hatred: the awakening of a former white nationalist**. New York: Doubleday, 2018.

SIMPSON, R. M. “Lost, enfeebled, and deprived of its vital effect”: Mill’s exaggerated view of the relation between conflict and vitality. **Aristotelian Society Supplementary Volume**, v. 95, n. 1, p. 97-114, jul. 2021. Disponível em: <https://doi.org/10.1093/arisup/akab006>. Acesso em: 1 mar. 2025.

STONE, R. The Daily transcript: interview with former white nationalist Derek Black. **The New York Times**, New York, 22 ago. 2017. Disponível em: <https://>

www.nytimes.com/2017/08/22/podcasts/the-daily-transcript-derek-black.html.
Acesso em: 1 mar. 2025.

SUNSTEIN, C. R. The law of group polarization. **Journal of Political Philosophy**, v. 10, n. 2, p. 175-195, 2002. Disponível em: <https://doi.org/10.1111/1467-9760.00148>. Acesso em: 1 mar. 2025.

TABER, C. S.; LODGE, M. Motivated skepticism in the evaluation of political beliefs. **American Journal of Political Science**, v. 50, n. 3, p. 755-769, jul. 2006. Disponível em: <http://www.jstor.org/stable/3694247>. Acesso em: 1 mar. 2025.

TALISSE, R. B. **Overdoing democracy: why we must put politics in its place**. New York: Oxford University Press, 2019.

VAN PROOIJEN, J.; SPADARO, G.; WANG, H. Suspicion of institutions: how distrust and conspiracy theories deteriorate social relationships. **Current Opinion in Psychology**, v. 43, p. 65-69, 2022. Disponível em: <https://doi.org/10.1016/j.copsyc.2021.06.013>. Acesso em: 1 mar. 2025.

WRIGHT, C. Wittgenstein on mathematical proof. **Royal Institute of Philosophy Supplement**, v. 28, p. 79-99, 1990. Disponível em: <https://doi.org/10.1017/S1358246100005257>. Acesso em: 1 mar. 2025.

YOUNG, I. M. **Inclusion and democracy**. Oxford: Oxford University Press, 2000.

ZOLLMAN, K. J. S. Network epistemology: communication in epistemic communities. **Philosophy Compass**, v. 8, p. 15-27, 2013. Disponível em: <https://doi.org/10.1111/j.1747-9991.2012.00534.x>. Acesso em: 1 mar. 2025.