



Revisão Narrativa

POR QUE MENTIRAS SE ESPALHAM MAIS RÁPIDO QUE A VERDADE: MECANISMOS COGNITIVOS E SOCIOTÉCNICOS DA DESINFORMAÇÃO NA ERA DIGITAL

WHY LIES SPREAD FASTER THAN TRUTH: COGNITIVE AND SOCIOTECHNICAL MECHANISMS OF MISINFORMATION IN THE DIGITAL AGE

Rogério F. Gonçalves¹

Rogério Ferreira Gonçalves **Afiliação:** Centro Universitário Evangélico de Goianésia (UNIEGO) **Autor correspondente:** rogerio.goncalves@docente.uniego.edu.br **Telefone:** 55 (62)984642280.

Resumo

A desinformação digital constitui um dos desafios mais prementes da sociedade contemporânea, com evidências robustas de que notícias falsas se difundem de forma significativamente mais rápida, profunda e ampla do que informações verdadeiras. O presente artigo, por meio de revisão narrativa da literatura, analisa os mecanismos cognitivos e sociotécnicos que explicam esse fenômeno, com ênfase em três eixos: o modelo de processamento dual de Kahneman, a hipótese spinozana de crença automática de Gilbert e os estudos empíricos sobre difusão diferencial de desinformação conduzidos por Vosoughi, Roy e Aral e por Pennycook e colaboradores. A literatura evidencia que a assimetria informacional não decorre prioritariamente da ação de robôs ou de vieses políticos motivados, mas sim da interação entre a arquitetura das plataformas digitais — que privilegiam a velocidade do impulso sobre a deliberação — e vieses cognitivos evolutivamente conservados, notadamente o viés de confirmação, a heurística da novidade e a tendência automática de acreditar antes de verificar. Os achados indicam que intervenções voltadas ao redirecionamento da atenção para a acurácia informacional, combinadas com estratégias de alfabetização midiática e letramento científico, constituem abordagens promissoras para mitigar o fenômeno. Conclui-se que o combate à desinformação exige compreensão integrada dos substratos psicológicos que tornam o cérebro humano vulnerável a conteúdos falsos, emocionalmente carregados e epistemicamente sedutores.

Palavras-chave: Desinformação; Desordem informacional; Viés de confirmação; Processamento dual; Redes sociais; Alfabetização midiática; Cognição; fake news.

Abstract

Digital misinformation constitutes one of the most pressing challenges of contemporary society, with robust evidence that false news spreads significantly faster, deeper, and more broadly than true information. This article, through a narrative literature review, analyzes the cognitive and sociotechnical mechanisms that explain this phenomenon, emphasizing three axes: Kahneman's dual-processing model, Gilbert's Spinozan hypothesis of automatic belief, and the empirical studies on the differential diffusion of misinformation conducted by Vosoughi, Roy, and Aral and by Pennycook and

colleagues. The literature shows that informational asymmetry does not stem primarily from robot action or motivated political biases, but rather from the interaction between the architecture of digital platforms — which privilege impulse speed over deliberation — and evolutionarily conserved cognitive biases, notably confirmation bias, the novelty heuristic, and the automatic tendency to believe before verifying. The findings indicate that interventions aimed at redirecting attention toward informational accuracy, combined with media literacy and scientific literacy strategies, constitute promising approaches to mitigate the phenomenon. It is concluded that combating misinformation requires an integrated understanding of the psychological substrates that make the human brain vulnerable to false, emotionally charged, and epistemically seductive content.

Keywords: Misinformation; Information disorder; Confirmation bias; Dual processing; Social media; Media literacy; Cognition; Fake news.

1. INTRODUÇÃO

Na noite de 30 de outubro de 1938, o jovem radialista norte-americano Orson Welles levou ao ar, pela rede CBS, uma adaptação do romance *A Guerra dos Mundos*, de H. G. Wells. A peça narrava, em formato de boletim jornalístico, uma invasão marciana em Nova Jersey. Embora Welles tenha advertido, no início da transmissão, que se tratava de ficção, milhares de ouvintes que sintonizaram após o aviso foram expostas a um locutor que descrevia, em tom de urgência, naves extraterrestres descendo sobre a Terra, exércitos sendo dizimados e cidades em chamas. O pânico foi documentado: telefonemas congestionaram centrais, famílias fugiram de casa, houve quem se trancasse no porão com armas (CANTRIL; GAUDET; HERZOG, 1940).

O episódio de *A Guerra dos Mundos* é frequentemente lembrado como curiosidade histórica. Contudo, o que ocorreu naquela noite constitui a prévia de um fenômeno que, oito décadas depois, se tornaria a arquitetura invisível da comunicação contemporânea: a desinformação se propaga com velocidade e eficiência dramaticamente superiores às da verdade, e o faz explorando vulnerabilidades cognitivas que antecedem em muito a existência da internet.

Em 2018, Vosoughi, Roy e Aral publicaram na revista *Science* um estudo que analisou aproximadamente 126 mil notícias verificadas, compartilhadas por cerca de 3 milhões de usuários do Twitter ao longo de 11 anos. O achado central foi contundente: a desinformação se difunde de forma significativamente mais distante, mais rápida, mais profunda e mais ampla do que a verdade em todas as categorias de informação analisadas, com efeitos particularmente pronunciados para notícias políticas (VOSOUGHI; ROY; ARAL, 2018).

Diante desse cenário, o presente artigo tem como objetivo analisar, por meio de revisão narrativa da literatura, os mecanismos cognitivos e sociotécnicos que explicam por que a desinformação se dissemina mais rapidamente do que a verdade, com ênfase em três eixos interdependentes: o modelo de processamento dual proposto por Kahneman (2011), a hipótese spinozana de crença automática formulada por Gilbert (1991) e as evidências empíricas sobre difusão diferencial de desinformação nas plataformas digitais.

2. MATERIAL E MÉTODOS

Trata-se de revisão narrativa da literatura, método indicado para a síntese de conhecimentos sobre temáticas amplas que demandam integração de evidências oriundas de múltiplas tradições disciplinares, sem a pretensão de exaustividade característica das revisões sistemáticas (ROTHER, 2007).

A busca bibliográfica foi conduzida entre junho e julho de 2026 nas bases de dados PubMed, Scopus, Web of Science, PsycINFO e Google Scholar. Utilizaram-se os seguintes descritores, em português e inglês, combinados por operadores booleanos: "misinformation", "disinformation", "fake news", "spread of false news", "cognitive bias", "dual process theory", "confirmation bias", "social media", "attention AND accuracy", "belief AND automatic processing", "information disorder", "desinformação", "desordem informacional", "viés de confirmação", "processamento dual". Não se aplicou restrição de data de publicação, considerando-se a necessidade de incluir obras seminais dos campos da psicologia cognitiva e da psicologia social.

Os critérios de inclusão abrangeram: (a) artigos originais com dados empíricos sobre difusão de desinformação em plataformas digitais; (b) estudos experimentais sobre mecanismos cognitivos de crença e compartilhamento de informações; (c) revisões sistemáticas e meta-análises sobre desinformação e vieses cognitivos; (d) obras teóricas seminais sobre processamento dual, crença automática e heurísticas cognitivas. Excluíram-

se publicações de caráter exclusivamente computacional sem discussão de implicações psicológicas, bem como artigos de opinião sem respaldo empírico.

Os achados foram organizados em quatro categorias temáticas para apresentação dos resultados: difusão diferencial da desinformação, processamento dual e compartilhamento impulsivo, a hipótese spinozana de crença automática, e intervenções baseadas em atenção à acurácia.

3. RESULTADOS

3.1 A difusão diferencial da desinformação: o estudo do MIT

Vosoughi, Roy e Aral (2018) analisaram a difusão de todas as notícias verificadas — classificadas como verdadeiras ou falsas por seis organizações independentes de checagem de fatos com 95% a 98% de concordância — distribuídas no Twitter entre 2006 e 2017. O conjunto de dados compreendeu aproximadamente 126 mil histórias tuitadas por cerca de 3 milhões de pessoas mais de 4,5 milhões de vezes.

Os resultados revelaram que a falsidade se difundiu de forma significativamente mais distante, mais rápida, mais profunda e mais ampla do que a verdade em todas as categorias de informação. A probabilidade de uma notícia falsa ser retuitada foi 70% superior à de uma notícia verdadeira. Enquanto as notícias verdadeiras raramente alcançavam mais de mil pessoas, as falsas frequentemente atingiam dezenas de milhares (VOSOUGHI; ROY; ARAL, 2018).

Contrariando a sabedoria convencional, os robôs (bots) aceleraram a difusão de notícias verdadeiras e falsas na mesma proporção, indicando que a desinformação se espalha mais do que a verdade porque seres humanos — e não robôs — são mais propensos a compartilhá-la (VOSOUGHI; ROY; ARAL, 2018).

A análise das emoções associadas às respostas revelou que notícias falsas inspiravam medo, repulsa e surpresa, enquanto notícias verdadeiras inspiravam antecipação, tristeza, alegria e confiança. Adicionalmente, as notícias falsas eram percebidas como significativamente mais novas (*novel*) do que as verdadeiras, sugerindo que os usuários compartilham desinformação, em parte, porque ela parece mais original e inesperada (VOSOUGHI; ROY; ARAL, 2018).

3.2 O cérebro que acredita primeiro e pergunta depois: processamento dual e compartilhamento impulsivo

Kahneman (2011), em sua obra seminal *Rápido e Devagar: Duas Formas de Pensar*, propôs um modelo de processamento dual que distingue dois sistemas cognitivos: o Sistema 1, rápido, intuitivo, emocional e automático, que opera com pouco ou nenhum esforço e sensação de controle voluntário; e o Sistema 2, lento, analítico, deliberado e lógico, que requer atenção e é acionado quando uma decisão demanda esforço cognitivo consciente. O Sistema 1 é eficiente na maioria das situações cotidianas, mas é suscetível a vieses sistemáticos e erros previsíveis, particularmente quando confrontado com informações emocionalmente carregadas (KAHNEMAN, 2011).

As plataformas de mídia social são arquitetadas para operar na velocidade do Sistema 1. A rolagem contínua de *feeds*, a sucessão rápida de estímulos e a efemeridade dos conteúdos desencorajam o acionamento do Sistema 2. Quando uma manchete emocionalmente impactante aparece, o Sistema 1 reage com medo, surpresa ou indignação, e a decisão de compartilhar ocorre antes que o Sistema 2 possa intervir com uma avaliação crítica de veracidade.

Pennycook e Rand (2019) testaram diretamente essa hipótese, investigando se a suscetibilidade a notícias falsas é melhor explicada por raciocínio motivado (as pessoas acreditam no que é politicamente conveniente) ou por falta de deliberação analítica (as pessoas acreditam em notícias falsas porque simplesmente não param para pensar). Em uma série de estudos com amostras representativas, os autores encontraram que a capacidade de discernir entre notícias verdadeiras e falsas estava positivamente correlacionada com medidas de reflexão cognitiva e negativamente correlacionada com a tendência a confiar em intuições. A conclusão foi que a suscetibilidade à desinformação é melhor explicada pela preguiça cognitiva do que pelo viés politicamente motivado (PENNYCOOK; RAND, 2019).

Em estudo subsequente, Bago, Rand e Pennycook (2020) utilizaram um paradigma de duas respostas (*two-response paradigm*) no qual os participantes primeiro emitiam um julgamento intuitivo sob pressão de tempo e carga de memória de trabalho, para depois reconsiderar a decisão sem restrições. Os resultados demonstraram que a deliberação corrigia erros intuitivos: os participantes acreditavam menos em manchetes falsas quando tinham oportunidade de deliberar, enquanto a crença em manchetes verdadeiras permanecia estável ou aumentava. Esse achado corrobora a hipótese clássica de que a falha em discernir entre informação verdadeira e falsa decorre primordialmente da ausência de deliberação, e não do uso da deliberação para racionalizar crenças ideologicamente convenientes (BAGO; RAND; PENNYCOOK, 2020).

3.3 A hipótese spinozana: compreender é acreditar

Gilbert (1991) propôs uma tese provocativa sobre a natureza do processo de crença, contrastando o modelo cartesiano — segundo o qual a compreensão de uma proposição e sua aceitação como verdadeira são operações psicológicas sequenciais e independentes — com o modelo spinozano — segundo o qual a aceitação de uma ideia é parte automática de sua compreensão, e a rejeição ocorre apenas posteriormente, por meio de um processo secundário e cognitivamente custoso.

Em uma série de experimentos que manipularam a disponibilidade de recursos cognitivos (por meio de interrupções, sobrecarga de memória e pressão de tempo), Gilbert (1991) demonstrou que a interrupção do processamento cognitivo tornava os participantes mais propensos a aceitar proposições falsas como verdadeiras, mas não os tornava mais propensos a rejeitar proposições verdadeiras. Em outras palavras, quando os recursos cognitivos são escassos, as pessoas tendem a acreditar no que leem, e não a duvidar.

A implicação para o ecossistema informacional contemporâneo é direta. As plataformas digitais, ao promoverem um consumo de informação fragmentado, acelerado e superficial, criam precisamente as condições de escassez cognitiva nas quais o mecanismo spinozano de crença automática opera com máxima eficiência. O usuário que rola rapidamente o *feed* está, na prática, com seus recursos de processamento do Sistema 2 comprometidos, tornando-se mais vulnerável à aceitação acrítica de informações falsas.

3.4 Para além do déficit: intervenções baseadas em atenção à acurácia

Pennycook et al. (2021), em estudo publicado na *Nature*, investigaram uma aparente contradição: a maioria das pessoas afirma valorizar a precisão das informações que compartilha, mas a veracidade das manchetes tem pouco efeito sobre as intenções de compartilhamento. Os autores demonstraram, em quatro experimentos de levantamento e um experimento de campo no Twitter, que induzir sutilmente os participantes a pensar sobre acurácia — por exemplo, solicitando que avaliassem a veracidade de uma única manchete não política antes de decidir quais conteúdos compartilhariam — aumentava

significativamente a qualidade das notícias subsequentemente compartilhadas (PENNYCOOK et al., 2021).

Esse achado é consistente com a hipótese da desatenção: as pessoas frequentemente compartilham desinformação não porque valorizam o partidarismo acima da verdade, mas porque sua atenção, no contexto das mídias sociais, está focada em fatores diferentes da acurácia. A intervenção de redirecionamento atencional é simples, escalável e passível de implementação pelas próprias plataformas (PENNYCOOK et al., 2021).

O **Quadro 1** sintetiza os principais mecanismos cognitivos de vulnerabilidade à desinformação identificados na literatura revisada.

Quadro 1 — Mecanismos cognitivos de vulnerabilidade à desinformação digital

Mecanismo	Descrição	Referência seminal
Crença automática (hipótese spinozana)	A aceitação de uma proposição como verdadeira é parte automática de sua compreensão; a rejeição requer esforço cognitivo adicional	Gilbert (1991)
Processamento Sistema 1*	Decisões de compartilhamento ocorrem por impulso intuitivo, sem acionamento do Sistema 2 analítico*	Kahneman (2011)
Heurística da novidade	Informações novas e inesperadas são mais compartilhadas independentemente de sua veracidade	Vosoughi, Roy e Aral (2018)
Viés de confirmação	Tendência a buscar, interpretar e lembrar informações que confirmam crenças preexistentes	Nickerson (1998)
Preguiça cognitiva*	A falha em discernir entre verdade e falsidade decorre mais da ausência de deliberação do que de raciocínio motivado	Pennycook e Rand (2019)
Emocionalidade negativa*	Conteúdos que evocam medo, repulsa e surpresa são compartilhados com maior frequência	Vosoughi, Roy e Aral (2018)

Fonte: Elaborado pelo autor com base na literatura revisada (2026).

DISCUSSÃO

Os resultados desta revisão convergem para a constatação de que a vantagem competitiva da desinformação sobre a verdade no ambiente digital não decorre de um fator isolado, mas de uma confluência de mecanismos cognitivos evolutivamente conservados com uma arquitetura sociotécnica que os explora de forma inédita na história da comunicação humana.

O modelo de processamento dual de Kahneman (2011) fornece a moldura teórica mais abrangente para compreender o fenômeno. As plataformas digitais foram projetadas para engajar o Sistema 1 — rápido, intuitivo, emocional — e desencorajar o acionamento do Sistema 2 — lento, analítico, deliberado. Como demonstraram Bago, Rand e Pennycook (2020), quando os participantes têm oportunidade de deliberar, sua capacidade de discernir entre informações verdadeiras e falsas melhora significativamente. O problema é que o *design* das plataformas — com *feeds* infinitos, rolagem contínua, estímulos audiovisuais intermitentes e recompensas sociais imediatas (*likes*, compartilhamentos) — cria um ambiente no qual a deliberação é a exceção, não a regra.

O achado de Vosoughi, Roy e Aral (2018) de que notícias falsas inspiram medo, repulsa e surpresa, enquanto notícias verdadeiras inspiram emoções de menor ativação (antecipação, tristeza, confiança), é particularmente revelador. Essas emoções de alta ativação são processadas pela amígdala e por circuitos subcorticais que operam em velocidades muito superiores às do córtex pré-frontal — a sede do pensamento analítico e do controle inibitório. A desinformação, portanto, explora uma via neural rápida que é evolutivamente anterior à capacidade de raciocínio abstrato.

A hipótese spinozana de Gilbert (1991) adiciona uma camada explicativa crucial: não apenas as plataformas favorecem o processamento intuitivo sobre o analítico, como o próprio ato de compreender uma proposição já envolve, automaticamente, sua aceitação provisória como verdadeira. A rejeição — a dúvida, o ceticismo, a verificação — é um processo secundário, energeticamente custoso e facilmente suprimido por condições de escassez cognitiva. Ora, o consumo de informação nas mídias sociais é caracterizado precisamente por fragmentação, velocidade e sobrecarga — condições ideais para que o mecanismo spinozano opere sem contrapeso.

Contrariando uma intuição amplamente difundida, os dados não sustentam a hipótese de que robôs ou agentes automatizados sejam os principais responsáveis pela difusão de desinformação. Vosoughi, Roy e Aral (2018) demonstraram que os bots aceleram a disseminação de informações verdadeiras e falsas na mesma proporção. O problema, portanto, não é primariamente tecnológico, mas humano: são as pessoas que compartilham desinformação, movidas por impulsos emocionais e vieses cognitivos que as plataformas aprenderam a explorar com precisão.

A implicação prática mais promissora dos estudos revisados é a de que intervenções relativamente simples — como solicitar que o usuário avalie a acurácia de uma informação antes de compartilhá-la — podem produzir efeitos significativos na qualidade do ecossistema informacional. Pennycook et al. (2021) demonstraram que esse tipo de intervenção, conhecida como *nudging* de acurácia, é eficaz tanto em ambiente controlado de laboratório quanto em experimento de campo no Twitter. Trata-se de uma abordagem que não depende de remoção de conteúdo, de verificação externa exaustiva ou de mudanças estruturais profundas no modelo de negócio das plataformas — embora estas possam ser desejáveis por outras razões.

Esta revisão apresenta limitações. Por seu caráter narrativo, não emprega protocolo sistemático de seleção e avaliação da qualidade dos estudos, o que restringe a reprodutibilidade. Adicionalmente, o campo da pesquisa sobre desinformação digital evolui rapidamente, e novos achados podem matizar ou complementar as conclusões aqui apresentadas. Estudos futuros deveriam investigar a eficácia comparativa de diferentes modalidades de intervenção — *nudges* de acurácia, alfabetização midiática, *prebunking* e *debunking* — em contextos culturais diversos, incluindo populações de países em desenvolvimento, onde a penetração das plataformas é elevada, mas os níveis de letramento digital são heterogêneos.

CONCLUSÕES

A presente revisão narrativa permite concluir que a disseminação assimétrica de desinformação nas plataformas digitais é sustentada por três pilares interdependentes: (a) a arquitetura das plataformas, que privilegia a velocidade do impulso sobre a deliberação, operando na faixa de processamento do Sistema 1; (b) o viés evolutivamente conservado de acreditar antes de verificar, formalizado na hipótese spinozana de crença automática; e (c) a vantagem epistêmica do conteúdo falso, que tende a ser mais novo, mais emocionalmente ativador e mais alinhado a vieses de confirmação do que o conteúdo verdadeiro.

As evidências indicam que intervenções de redirecionamento atencional voltadas à acurácia — em combinação com programas de alfabetização midiática e letramento científico — constituem a estratégia mais promissora e empiricamente fundamentada para mitigar o fenômeno, superando abordagens baseadas exclusivamente em remoção de conteúdo ou em correção reativa (*debunking*).

A compreensão dos mecanismos psicológicos que tornam o cérebro humano vulnerável à desinformação não é apenas uma questão acadêmica: é um pré-requisito para

o desenho de políticas públicas, práticas educacionais e arquiteturas tecnológicas que preservem os benefícios da conectividade digital sem sacrificar a capacidade coletiva de distinguir fato de ficção.

Declaração de conflito de interesse: O autor declara não haver conflito de interesse de natureza financeira, institucional ou pessoal que possa influenciar a interpretação dos resultados apresentados.

Abreviações: FOMO — *Fear of Missing Out*; MIT — Massachusetts Institute of Technology.

REFERÊNCIAS

- Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, 149(8), 1608–1613. <https://doi.org/10.1037/xge0000729>
- Cantril, H., Gaudet, H., & Herzog, H. (1940). *The invasion from Mars: A study in the psychology of panic*. Princeton University Press.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46(2), 107–119. <https://doi.org/10.1037/0003-066X.46.2.107>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50. <https://doi.org/10.1016/j.cognition.2018.06.011>
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855), 590–595. <https://doi.org/10.1038/s41586-021-03344-2>
- Rother, E. T. (2007). Revisão sistemática × revisão narrativa. *Acta Paulista de Enfermagem*, 20(2), v–vi. <https://doi.org/10.1590/S0103-21002007000200001>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>

