

UMA ABORDAGEM SEMIÓTICO-COGNITIVA DO FENÔMENO DA DESINFORMAÇÃO

A semiotic-cognitive approach to the phenomenon of disinformation

Una aproximación semiótico-cognitiva al fenómeno de la desinformación

Gustavo Rick Amaral¹
Tarcísio Cardoso²

DOI: 10.31501/esf.v1i32.15505

Resumo: Neste artigo, pretendemos reenquadrar o conceito de desinformação a partir de Bayes e Peirce, de modo a designar um desvio ou mesmo bloqueio em relação a um processo informativo ideal. Defendemos que embora grande parte dos fenômenos agrupados sob o rótulo desinformação sejam categorizados genericamente como mentira ou enganação, os fatores diferenciais que justificam o neologismo dizem respeito a condições particulares da mídia digital, que vamos denominar sistematicidade e escalabilidade.

Palavras-chaves: Desinformação. Atualização informativa. Mídia digital. Sistematicidade. Escalabilidade.

Abstract: In this article, we reexamine the concept of disinformation through the lenses of Bayes and Peirce, viewing it as a deviation or blockage in an ideal information process. We argue that while specific instances of disinformation are generally seen as lies or deception, the distinguishing factors that underscore the novelty of the term relate specifically to the conditions of digital media. We refer to these conditions as systematicity and scalability.

Keywords: Disinformation. Information update. Digital media. Systematicity. Scalability.

Resumen: En este artículo pretendemos replantear el concepto de desinformación basándose en Bayes y Peirce, para designar una desviación o incluso un bloqueo en relación con un proceso de información ideal. Sostenemos que si bien gran parte de los fenómenos agrupados bajo la etiqueta de desinformación se categorizan genéricamente como mentira o engaño, los factores diferenciadores que justifican el neologismo se relacionan con condiciones particulares de los medios digitales, a las que llamaremos sistematicidad y escalabilidad.

Palabras-clave: Desinformación. Actualización de información. Medios digitales. Sistematicidad. Escalabilidad.

¹ Doutor; Universidade Anhembi-Morumbi, São Paulo - SP, Brasil. gustrick@gmail.com; <https://orcid.org/0000-0002-0063-6119>.

² Doutor; Universidade Federal da Bahia, Salvador - BA, Brasil. tcardoso@ufba.br; <https://orcid.org/0000-0002-1093-5307>.

Esta pesquisa teve apoio da FAPESB (Fundação de Amparo à Pesquisa da Bahia).

1. Introdução

A concepção de desinformação mais adotada pela comunidade científica se baseia na noção de desordem informacional de Derakhshan e Wardle (2017), que compreende três termos em inglês: *disinformation*, *misinformation* e *malinformation*. A palavra em português “desinformação” tem sido usada para designar todas as nuances da desordem informacional, abarcando uma diversidade tão grande de fenômenos e definições tão hesitantes, por vezes excessivamente metafóricas e com contornos vagos, que é o caso de se perguntar: este neologismo com ares de oximoro se justifica? Tem alguma utilidade para análises acadêmicas e científicas mais profundas? Nossa resposta para ambas as questões é positiva, mas difere da maioria dos estudos que tentam definir desinformação (Bennett & Livingston, 2018; Fallis, 2015; Oliveira, 2020; Santaella, 2023; Wardle & Derakhshan, 2017).

Embora grande parte dos fenômenos que, nos últimos anos, passamos a agrupar sob o rótulo de “desinformação” – *fake news*, *deep fakes*, negacionismos, pseudociência, etc. – sejam discursos falsos que poderiam, portanto, ser categorizados genericamente como “mentira” ou “enganação”, o que justifica, a nosso ver, o neologismo é que tais discursos podem, graças a condições particulares do meios de comunicação digitais dadas por alguns desenvolvimentos técnicos como filtros algorítmicos turbinados por IA e estratégias de *microtargeting*, ser produzidos de forma sistemática e numa escala sem precedentes, o que vamos denominar respectivamente *sistematicidade* e *escalabilidade*. No modelo teórico que pretendemos apresentar neste artigo convergimos aparatos conceituais dos estudos da cognição com os do campo da comunicação e da semiótica para dar tratamento ao fenômeno da desinformação a partir de três níveis: *individual*, *social imediato*, *social amplo*.

No modelo teórico de base semiótico-comunicacional que estamos introduzindo neste artigo, a desinformação deve ser entendida como um descaminho em relação ao processo informativo *normal*. O termo “normal” comparece aqui no sentido de padrão normativo, uma idealização racional do modo pelo qual a mente humana *deveria* usar informação para formar conhecimento ou fixar crença sobre determinado tópico. Tem-se, assim, uma expectativa projetada pelo padrão e um desvio do curso informativo da norma. Assim, na nossa concepção, desinformação é um termo que serve para designar a medida desse afastamento em relação ao padrão normativo.

Nossa definição de desinformação, à primeira vista, é contraintuitiva, pois a entendemos como um tipo de estratégia com aspectos comunicacionais e semióticos que desencaminha a atualização informativa com o objetivo de proteger uma crença particular, que denominamos *crença de fundo*³. Com isso nos afastamos da maioria das abordagens que a definem como sinônimo de mentira ou representação falsa e ainda de outras que a entendem como qualquer fator que atrapalhe o fluxo comunicacional (Bennett & Livingston, 2018; Fallis, 2015; Wardle & Derakhshan, 2017). A nosso ver, as acepções mais comuns consideram desinformação quase como um sinônimo de mentira ou ruído (conforme conceituado na teoria da informação de Shannon [cf. Shannon & Weaver, 1975]). Denominemos esta acepção de “desinformação como ruído” de *sentido lato* de desinformação. Em contraposição, denominemos nossa proposta de *sentido estrito*.

³ É possível, para fins de análise, diferenciar a desinformação como designando, por um lado, estratégias gerais que operam com a finalidade de desencaminhamento e designando, por outro lado, aplicações concretas (instâncias) dessas estratégias.

Definição de desinformação *strictu sensu*: A desinformação é um conjunto composto por representações (essencialmente) falsas que, de forma sistemática e escalável, desvia do seu curso normal (e, no limite, bloqueia) o processo de atualização informativa referente a uma crença de fundo.

Reparemos que desinformação neste sentido estrito não se reduz a uma representação falsa (ou mentira, que é um caso particular de representação falsa), mas constitui um conjunto de representações que tem (de modo intencional ou não) um determinado efeito de desencaminhar o curso da semiose⁴: opera um desvio em relação ao curso normal do processo de atualização informativa. O alvo da estratégia de desinformação é o que estamos chamando de crença de fundo. O objetivo é protegê-la. Denomina-se "crença de fundo", porque a estratégia opera de forma indireta para atingir esse objetivo. Ela mobiliza representações (essencialmente) falsas para ir aos poucos desviando a semiose de seu "curso normal" (que é o da atualização informativa). Essas representações podem ter como objeto tópicos distintos da crença de fundo a ser protegida, podem a ela estar ligadas de forma indireta⁵. Com o desvio (e eventual bloqueio) ocasionado por essas representações atinge-se a meta de proteger a crença de fundo. Nas próximas três seções que compõem este artigo nos dedicaremos a explicar não apenas o que significa este "curso normal da atualização informativa" como também esses critérios de *sistematicidade* e *escalabilidade* que distinguem o que estamos chamando de

⁴ O conceito de semiose será explicado mais adiante neste texto.

⁵ Num exemplo desenvolvido na última seção deste artigo, solicitamos ao leitor para imaginar a seguinte situação: uma estratégia desinformativa que mobilize representações falsas associando uma minoria social a vícios, condutas moralmente reprováveis e até mesmo a crimes com o objetivo de reforçar uma crença de fundo segundo a qual esta mesma minoria seria a causa de um problema econômico pelo qual passa a sociedade como um todo. Reparemos que a crença de fundo diz respeito a um problema econômico, e as representações falsas que compõem a estratégia desinformativa têm como objeto práticas ou ações relativamente distantes do campo econômico. Mesmo as representações estando desligadas logicamente da crença de fundo, a força das imagens criadas pelas representações falsas pode – de modo eficiente – induzir o processo interpretativo de muitos receptores não apenas a reforçar o conteúdo desta crença como também bloquear (interpretativamente) informações contrárias a ela.

desinformação no sentido estrito de outros fenômenos próximos, como representação falsa, mentira, ruído, etc.

2. Atualização informativa: o impacto da informação numa crença

Na definição apresentada acima de desinformação (*stricto sensu*), introduzimos o conceito de *atualização informativa* para designar uma *forma idealizada* pela qual uma mente ou sistema cognitivo deveria lidar com informação com o objetivo de construir conhecimento. Nesta seção, explicamos como tal idealização de um processo informativo pode ser capturada a partir de aparatos teórico-conceituais tão distintos como a estatística bayesiana e a semiótica peirceana.

O teorema de Bayes é uma fórmula para lidar com a probabilidade condicional de um evento. Probabilidade condicional é a ideia de se calcular a probabilidade de que algo ocorra com base em determinadas condições; por exemplo, a probabilidade de ocorrer *A* dado que ocorreu *B*. A forma mais simples desse teorema mede o impacto que uma nova evidência (uma informação) tem sobre as chances de uma crença que temos em algo (um argumento ou uma teoria específica, por exemplo) se provar correta (Matthews, 2017, p. 152). Não vamos descer nos detalhes da fórmula matemática. É suficiente reter que ela nos permite calcular o grau de convicção que *deveríamos ter* em uma crença a partir do momento que nos chega uma informação nova (evidência). Bayes nos fornece uma forma de enxergar como funcionaria uma máquina ideal, racional de atualização de crenças.

Entre as aplicações mais notáveis da lógica de “atualização informativa” de Bayes está o Sistema de Planejamento Otimizado de Busca e Salvamento, SAROPS (sigla em inglês para *Search and Rescue*

Optimal Planning System). Dois célebres casos em que o SAROPS teve uma aplicação muito bem bem-sucedida foi para encontrar a caixa-preta do avião da AirFrance que caiu no oceano Atlântico em 2009 e para resgatar, em 2013, um pescador que caiu do seu barco e, usando suas botas como salva-vidas, ficou 12 horas em mar aberto a 60 km da costa de Long Island (EUA).

Vejamos no caso do resgate do pescador como opera o SAROPS (Matthews, 2017, p. 146). Uma vez alimentada com uma série de dados – por exemplo, sobre as condições das correntes marítimas naquele momento, sobre as condições de visibilidade, sobre o último local onde o homem foi visto etc. – a máquina gera um mapa de probabilidades da área de busca. É uma espécie de “mapa de calor” dividido em células coloridas. De posse desse mapa, as equipes de busca tinham um guia muito eficiente orientando suas ações em campo. Com um helicóptero, elas sobrevoavam os setores com maior probabilidade de encontrar o homem perdido e, quando não o encontravam, alimentavam a máquina com essa informação.

O segredo do SAROPS é que, quando chega a informação de que o objeto da busca não foi encontrado num determinado setor, o mapa de probabilidades é atualizado e a sua distribuição de cores muda. O teorema de Bayes é utilizado justamente para essa atualização informativa, que é feita de acordo com um modelo no qual estão fixados os modos pelos quais cada tipo de informação que chega é interpretada. No âmbito do modelo, não há *variabilidade interpretativa*. O modelo bayesiano que está subjacente à máquina interpreta os dados sempre da mesma forma. O peso dado a cada evidência que chega como *input* só muda se o modelo for alterado.

A forma como o SAROPS lida com a informação e atualiza a crença pode ser considerada um *modo idealizado* se compararmos com a maneira pela qual opera a cognição humana, porque o

processo interpretativo que a máquina faz do *input*/dado que chega do ambiente leva em conta um número limitado de parâmetros que estão *abstraídos* de uma diversidade de fatores que costumam intervir no “mundo mental” humano.

A cognição humana – pelo relato que temos de nossa melhor ciência de base empírico-experimental⁶ – está continuamente submetida a interferência de vieses que podem ser ocasionados tanto por “tendências naturais” (adquiridas pela trajetória evolutiva da espécie) como por crenças e conceitos estabelecidos em tempos históricos e dependentes de fatores culturais e sociais. Diferentemente da cognição humana, quando informamos uma máquina que o objeto das buscas não foi encontrado em determinado setor, ela simplesmente absorve essa informação de acordo com o peso de evidência que tem fixado em seu modelo. Por exemplo, é impensável a situação em que a máquina adquirisse uma “postura cética” (em relação à informação apresentada) por *acreditar* que o agente de busca que estava no helicóptero era incompetente ou por ter uma *concepção estereotipada* a respeito da etnia ou da cor da pele do agente. Máquinas como o SAROPS não têm, neste sentido, vieses ou concepções prévias que possam interferir no processamento de uma informação.

Já a mente humana – mesmo as treinadas para isso – não consegue julgar informação dessa forma idealizada, neutra, desligada de crenças e concepções prévias. Como veremos, o que distingue o fenômeno da desinformação é justamente a alteração do peso de uma evidência, a modificação do peso de uma informação referente a uma crença de fundo. Essa alteração, que ocorre durante o processo interpretativo (na mente do receptor/intérprete), é favorecida por determinadas características

⁶ Para a “abordagem da racionalidade limitada” (*bounded rationality*) desenvolvida por Hebert Simon, cf. Simon, 1955 e 1982; para as abordagens dos vieses e heurísticas e economia comportamental, cf. Kahneman & Tversky, 1979; Thaler, 1980; Kahneman, 2003 e 2011; Sunstein, 2020; Thaler & Sunstein, 2020; Kahneman, Sibony & Sunstein, 2021.

presentes tanto na mensagem/signo que faz parte da estratégia desinformativa como no contexto e na capacidade interpretativa do receptor/intérprete.

Já no campo da semiótica, especificamente naquele ramo desenvolvido a partir das propostas teóricas de C. S. Peirce, há um conceito que se aproxima dessa ideia geral de “atualização informativa” que estamos introduzindo neste artigo. É o conceito de *semiose*⁷, que designa o processo pelo qual um signo é interpretado em um outro signo (denominado interpretante) mais desenvolvido e tende, no longo prazo, a ir aproximando a representação do objeto representado. Embora a semiótica seja, na concepção peirceana, uma ciência normativa que descreva a *semiose* como um processo idealizado que tende para uma espécie de “estágio final” de representação (cujo termo técnico é interpretante final), Peirce tinha plena noção das vicissitudes e retrocessos de processos interpretativos/representativos reais operados por mentes corporificadas em organismos com uma história evolutiva, cultural, social, etc. e uma experiência individual particular. Por este motivo, a filosofia peirceana concebe o sujeito epistemológico como uma entidade coletiva, a comunidade indefinida (no tempo e no espaço) de intérpretes (cf. conceito de “comunidade cognitiva”, de Haack, 1982).

3. As práticas epistêmicas cotidianas e as funções semiótico-comunicacionais como pontos focais das estratégias de desinformação

O leitor deve ter reparado que estamos definindo o conceito de desinformação como o processo inverso ao da construção de conhecimento. Neste, utiliza-se o que podemos chamar de *estratégia*

⁷ Semiose é um conceito peirceano que quer se refere à ação triádica do signo. Peirce diz: “por ‘semiose’ entendo [...] uma ação ou influência que envolve uma cooperação entre três sujeitos, tais como o signo, seu objeto e seu interpretante” (EP2: 411 [1907]). Para Santaella (2008, p. 18), signo é um “esquema analítico elementar de um processo de continuidade que tanto regride quanto se prolonga ao infinito”, processo este entendido como semiose.

informativa para construir uma representação progressivamente verdadeira do mundo. Neste caso, o método, tal como Peirce (1877) argumenta, se baseia no uso de informação – acumulando evidências do modo mais controlado, imparcial e objetivo possível – para construir conhecimento. Em oposição, o que denominamos *estratégia desinformativa* tem a finalidade de proteger uma crença previamente estabelecida. Como essa proteção é operada em detrimento da construção de uma representação progressivamente verdadeira do mundo, então só podemos entendê-la, por contraste, como *manutenção da ignorância*.

Neste artigo, sustentamos que se deve analisar o modo como uma estratégia desinformativa funciona a partir de duas perspectivas complementares: a primeira delas é o ponto de vista do produtor da mensagem/signo; a segunda é o ponto de vista do receptor/intérprete. De um lado, tem-se uma estratégia comunicacional com finalidade desinformativa. Do outro, uma estratégia interpretativa também com finalidade desinformativa. Como deverá ficar claro ao longo de nossa exposição, essas são as duas faces da mesma moeda. Uma estratégia desinformativa pode ser definida, por um lado, como um tipo de ação comunicacional por parte do produtor da mensagem e, pelo outro lado, um tipo de processo interpretativo por parte do intérprete que – em ambos os casos – contribuem para o mesmo fim: evitar que uma crença prévia (na mente do intérprete) seja desmontada por evidências contrárias ou seja assaltada por dúvidas.

Nesta exposição de nosso modelo teórico, não vamos discutir o problema da intencionalidade por parte do produtor da mensagem ou por parte do intérprete. Há uma vasta discussão na literatura nacional (Oliveira, 2020; Recuero, 2020; Santaella, 2023) sobre o termo desinformação, em português, corresponder ao que, em inglês, se chama *disinformation*, *misinformation* ou ambos, e o único

consenso está no fato de que o primeiro inclui o paradigma da intencionalidade, já conhecido na comunidade internacional. Neste artigo, contudo, consideramos desinformação, inclusive não intencional, tanto nas camadas de emissão quanto nas de recepção de uma mensagem.

Da forma como estamos definindo o conceito, o objetivo geral da desinformação é desencaminhar a semiose para proteger uma crença prévia – que não é necessariamente falsa – denominada *crença de fundo*. As estratégias desinformativas são compostas por um certo número de *mensagens desinformativas* que constituem o meio para se atingir esse objetivo. Se abandonarmos uma visão idealizada a respeito da racionalidade humana e aceitarmos perspectivas mais descritivas, como a oferecida pela abordagem dos vieses e heurísticas, notaremos que esse objetivo não é tão difícil de ser atingido. Na verdade, o funcionamento mais comum da cognição humana parece ser o de proteger crenças previamente estabelecidas. O excepcional é submetê-las a algum tipo de escrutínio sistemático (como almejam filósofos e cientistas)⁸.

Se aceitarmos que as *práticas epistêmicas cotidianas* seguem na direção de proteção de crenças previamente estabelecidas, sobretudo aquelas mais centrais para o indivíduo (geralmente associadas à identidade individual, social, cultural, etc.), então notaremos que essa condição geral cria uma série de vulnerabilidades a serem exploradas por estratégias desinformativas. A eficiência de uma estratégia mais geral de desinformação ou de uma mensagem desinformativa em particular parece estar associada à capacidade de explorar tais fragilidades. Por esse motivo, elas são construídas tendo como pontos focais certos aspectos tanto do processo comunicacional como do processo interpretativo.

⁸ Nos próximos passos da pesquisa, procuraremos explicar – a partir do aparato conceitual elaborado por Peirce (2008, p. 33 – 58) no artigo “a fixação da crença” – como as *práticas epistêmicas* cotidianas se afastam das exigências de escrutínio sistemático do pensamento crítico.

Os pontos focais das estratégias de desinformação são o que vamos chamar de *funções semiótico-comunicacionais básicas*: a) captura da atenção; b) manutenção da atenção/engajamento; c) orientação do processo interpretativo; d) formação, uso ou fixação de conceitos (também chamados de “estruturas conceituais”); e) formação, uso ou fixação de crenças. Estas são denominadas funções básicas, porque pressupõe-se que devam ser, em alguma medida, cumpridas por todo processo comunicacional pelo lado do emissor e todo processo interpretativo (semiose) pelo lado do receptor. De um ponto de vista comunicacional, o emissor deve reconhecer que, para uma mensagem ser bem-sucedida, ela deve inicialmente chamar atenção do receptor, manter a atenção por um tempo (funções *a* e *b*) e ser capaz de, com auxílio do contexto, orientar o processo interpretativo (função *c*). A comunicação só se efetiva quando o receptor desenvolve um processo interpretativo (que “decodifica”, “acessa o conteúdo”) da mensagem e, para isso, mobiliza seu repertório, conceitos e crenças previamente formados (funções *d* e *e*) para aplicar regras interpretativas para a situação concreta. Como processos comunicacionais também têm o efeito de reforçar (i.e., fixar) conceitos e crenças prévias ou mesmo formar conceitos e crenças novas, subdividimos as funções *d* e *e* em três possibilidades (cf. Figura 1).

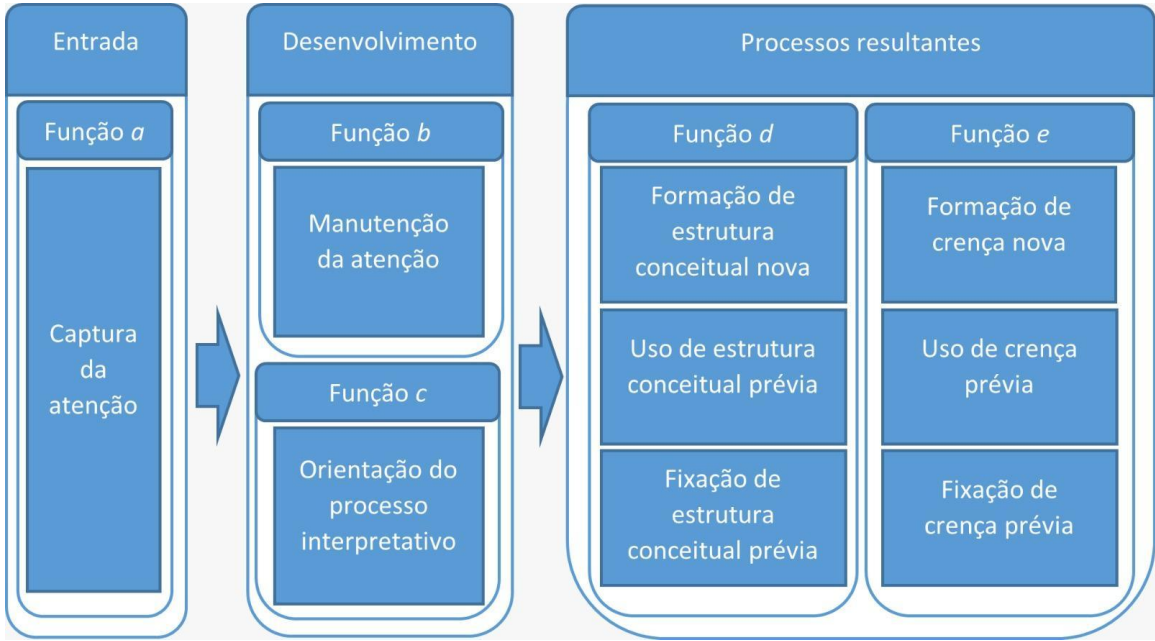


Figura 1 – Diagrama das funções semiótico-comunicacionais das práticas epistêmicas cotidianas (Fonte: elaborado pelo autor)

Como a estratégia desinformativa pretende desviar o processo interpretativo (de uma mensagem) do curso normal da atualização informativa de uma crença de fundo (ou, no limite, bloquear totalmente a atualização) justamente como o intuito de proteger a crença, as mensagens que compõem uma estratégia desinformativa procuram cumprir as funções acima descritas com vistas a essa finalidade de proteção. Em resumo, as estratégias desinformativas são determinadas formas de montar mensagens que acabam por favorecer o desvio ou mesmo bloqueio descrito na definição de desinformação. Explicadas as funções semiótico-comunicacionais, já estamos em condições de fecharmos o foco de

nossa exposição nos últimos dois conceitos da definição de desinformação *strictu sensu* introduzida acima: sistematicidade e escalabilidade.

4. Sistematicidade: desinformação não opera de forma isolada

A sistematicidade é um dos critérios que mobilizamos para distinguir desinformação *stricto sensu* de “fenômenos próximos” (como ruído, representação falsa, mentira, etc.). A sistematicidade que caracteriza a desinformação é um modo integrado e eficiente – ainda que epistêmica e moralmente questionável – de cumprir as funções semiótico-comunicacionais com o objetivo de desviar o processo interpretativo do seu curso normal de atualização informativa (e, no limite, bloqueá-lo).

Para nos aprofundarmos nessa conceituação de sistematicidade, voltemos à comparação entre mentira e desinformação. A diferença básica entre uma mentira e uma mensagem desinformativa é que, para a primeira, não há exigência de sistematicidade no sentido em que estamos usando este termo. Usa-se a palavra “mentira” geralmente para designar um ato comunicacional intencionado a criar uma representação falsa da realidade na mente do indivíduo-receptor. Já para haver desinformação, isso não é suficiente. Ela é uma representação que leva a interpretações enganosas sobre o estado de coisas, mas vai além, uma vez que é desenvolvida para proteger de forma contínua e sistemática uma crença de fundo.

Para apontarmos onde os conceitos de mentira e desinformação andam juntos e onde se separam em relação ao critério de sistematicidade, vamos a um exemplo prosaico e ingênuo: o proverbial caso da criança que mente que tomou banho. Nesse caso, a intenção em se criar uma

representação falsa parece ser mais imediata e episódica: apenas escapar do banho enganando os pais por um dia. O objetivo último seria simplesmente criar na mente dos pais uma representação falsa a respeito do *status* de sua higiene pessoal referente àquele curto período de tempo. Mas não poderíamos classificar esta mentira como um caso de desinformação? Poder-se-ia “forçar” o exemplo com a afirmação de que a criança, caso passasse a mentir todos os dias, teria a intenção não apenas de criar, mas também reforçar e proteger uma crença de fundo na mente dos pais, a crença segundo a qual ela tomaria banho todos os dias.

Vamos supor o cenário em que o infante realmente fosse vocacionado para arte de enganar e tivesse investido muito tempo desenvolvendo procedimentos que, ao serem repetidos estrategicamente ao longo do tempo, manteria os pais acreditando que ele tomaria banho diariamente. Reparemos que, para criar e proteger a crença de fundo na mente dos pais, a criança teria que não apenas executar os procedimentos, mas também se preocupar em cumprir o que chamamos acima de funções semiótico-comunicacionais. Não basta manipular evidências, por exemplo, molhando o cabelo e jogando fora um pouco do *shampoo* a cada dia. Ela deve fazer essas evidências falsas chegarem aos pais. As mensagens desinformativas devem atingir o alvo. Esse tipo de ação cumpre a função de “captura de atenção”, que, como veremos, é essencial para qualquer estratégia de desinformação ser bem-sucedida. Boas estratégias de desinformação se aproveitam de oportunidades para a crença de fundo ser reforçada.

Nesse cenário, é verdade que a mentira contada pela criança cumpriria o critério de sistematicidade. Entretanto, essa sistematicidade está longe de poder ser considerada eficiente. As mentiras contadas diariamente para os pais através dessa diversidade de estratagemas são difíceis de

serem mantidas. Não é apenas o custo cognitivo de mentir e o esforço para disfarçar a voz e adequar as expressões faciais durante a mentira. É também o tempo e energia investidos na preparação e execução de todos os procedimentos. Tudo nesse cenário é muito custoso para o mentiroso. Tanto é assim que o bom senso o aconselharia a abandonar a mentira. Dá muito mais trabalho mantê-la do que tomar o banho de verdade. Embora seja possível ao mentiroso cumprir as funções semiótico-comunicacionais de forma sistemática e até tenha êxito em proteger a crença de fundo (falsa) na mente dos pais, não se pode dizer que o faça de forma eficiente. Essa seria uma sistematicidade fraca. Por isso, o alto custo dos procedimentos.

A sistematicidade presente na desinformação é um modo *eficiente* de cumprir as funções semiótico-comunicacionais, dada a finalidade de desviar o curso da atualização informativa e proteger a crença de fundo. Assim, ainda que admitíssemos algum grau de sistematicidade, o exemplo acima de mentira sistemática do banho não se enquadraria como desinformação, pois opera num âmbito muito restrito, apenas uma família, a criança e seus pais, cujos processos comunicacionais são obviamente de pequena escala. Já as estratégias de desinformação operam com variáveis que estão além do ambiente imediato do indivíduo, o nível micro de seus grupos primários (como família, vizinhança, ambiente de trabalho, etc.), e passam a envolver variáveis que operam num nível macro, mobilizando fatores políticos, econômicos, sociais e culturais mais amplos e também os meios de comunicação de larga escala.

5. Escalabilidade: onde o monstro ganha corpo

A definição de escalabilidade que estamos mobilizando como critério de distinção para o fenômeno da desinformação *stricto sensu* é a capacidade do processo ganhar escala, aumentar seu alcance com baixo custo e com métodos acessíveis. Em escalas mais amplas de circulação de mensagens, como as permitidas pelos meios de comunicação de massa e pelos atuais meios digitais, entram em cena mecanismos de reforço que dão mais sustentabilidade à crença de fundo e – ponto que nos interessa mais de perto neste artigo – aos métodos de bloqueio da atualização informativa ou de desvio de seu curso normal. Compreender como operam esses mecanismos de reforço é essencial para o estudo do fenômeno da desinformação.

Os meios de comunicação estão diretamente relacionados à capacidade que uma sociedade tem de estabelecer um sistema de crenças, compartilhá-lo e reproduzi-lo ao longo do tempo. E escala faz diferença, pois a comunicação de larga escala – que é aquela permitida por livros ou jornais de ampla circulação, rádio, cinema, televisão e mais recentemente os meios digitais⁹ – implica que o material informativo que circula numa sociedade alcança uma audiência larga em um tempo curto, o que, por sua vez, eleva a probabilidade dos indivíduos serem expostos a informações favoráveis e contrárias ao que acreditam.

A comunicação de larga escala abre caminho para uma forma vertical de reforço do conteúdo de uma crença compartilhada que passa a operar retroalimentando a forma horizontal, que é aquela baseada na comunicação próxima, no contato direto entre as pessoas no ambiente social imediato. O

⁹ Apesar da diferença elementar entre a lógica de difusão centralizadas dos meios massivos e a lógica descentralizada (de compartilhamento) dos meios digitais, estamos considerando neste artigo ambos como comunicação de larga escala pela capacidade de atingir uma grande quantidade de receptores espalhados num vasto território num curto intervalo de tempo.

fator diferencial introduzido pelos meios digitais é a notável eficiência das ferramentas algorítmicas e técnicas de *microtargeting* disponíveis para o emissor produzir mensagens relevantes (Gillespie, 2014; Pariser, 2012) especificamente voltadas para proteger crenças de fundo em uma audiência.

Consideremos que informações referentes a uma determinada crença possam chegar para um indivíduo a partir de duas esferas de circulação de mensagens: a primeira delas é composta pelos meios de comunicação de larga escala; e a segunda, pelo seu ambiente social imediato. Lembrando que não é suficiente que uma mensagem contendo informação sobre uma crença circule no ambiente social, mas ela deve chegar no indivíduo, capturar sua atenção, ser “lida” e interpretada. É somente a partir do processo interpretativo que podemos avaliar a influência que a representação nova pode ter sobre uma crença de fundo previamente estabelecida na mente do indivíduo. Entretanto, ainda que a análise do processo interpretativo seja fundamental, ela não é suficiente. Embora o momento crucial – conforme antecipamos na seção anterior – do processo de desinformação seja a alteração do peso de uma evidência/informação referente a uma crença de fundo que ocorre na mente do receptor/intérprete, só podemos analisá-lo levando em conta as forças e os mecanismos de retroalimentação que entram em cena na escala micro e macro de circulação de mensagens.

Neste modelo teórico semiótico-cognitivo que introduzimos neste artigo, o fenômeno deve ser analisado em três níveis: o primeiro deles é o *nível individual*, que é o próprio receptor/intérprete das mensagens/signos da estratégia desinformativa; o segundo é o *nível social imediato* (do receptor/intérprete); e o terceiro é o nível macro, a sociedade como todo. Apenas com uma análise conjugada de processos nesses três níveis é possível capturar os fatores de sistematicidade e escalabilidade que julgamos centrais para o fenômeno da desinformação.

Em relação às estratégias desinformativas, o aumento de escalabilidade foi condicionado por avanços técnicos. No século passado, em que reinavam soberanas as mídias de massa, as técnicas mais refinadas de pesquisa de audiência eram consideravelmente complexas e exigiam mão de obra especializada, o que aumentava o custo econômico necessário para realizá-las. Atualmente, a capacidade de rastrear efeitos da comunicação no público está atualmente ultrapassando os limites do imaginável (cf. Bond *et al.*, 2012; Matz *et al.*, 2017). Para o presente artigo, o que importa é o modo pelo qual fatores tecnológicos, como algoritmos automatizados de seleção/recomendação de conteúdo (Gillespie, 2014), cria condições para se cumprirem as funções semiótico-comunicacionais de formas mais eficientes. O aumento no grau de refinamento da representação (das preferências) do público feita pelo emissor resulta numa eficiência maior para executar as funções semiótico-comunicacionais, sobretudo a da captura de atenção e a de engajamento (funções *a* e *b*, cf. na seção 2). A capacidade dos meios digitais de capturar atenção e engajar é tão grande que pesquisas vêm apontando efeitos negativos consideráveis no desenvolvimento cognitivo de crianças e jovens e também impactos na saúde mental da população em geral (cf. Desmurget, 2021; Haidt, 2024). Análise de dados e estratégias de *microtargeting* já demonstraram ter efeitos consideráveis em processos eleitorais e votações. Um caso de *desinformação customizada* que pode ser considerado paradigmático foi o que ficou conhecido como escândalo do *Cambridge Analytica* (Kaiser, 2019).

O segredo para o aumento de eficiência está na capacidade dos meios digitais rastrearem de forma automatizada (por isso, barata) os efeitos do processo comunicacional. O “pulo do gato” parece estar na automatização das diversas rotinas que compõem a representação das preferências do público. Esse tipo de automatização faz com que as funções semiótico-comunicacionais não apenas

sejam executadas de forma eficiente e barata, mas com uma característica denominada escalabilidade. No caso da desinformação, aqui é que o monstro realmente toma corpo. Para fecharmos nossa exposição, desenvolvamos um exemplo que – justamente pela presença das características de *sistematicidade* e *escalabilidade* – constitui um cenário, em termos de desinformação, diametralmente oposto ao da criança.

Vamos supor que um agente político absolutamente inescrupuloso tenha se candidatado para governar um país que passa por uma profunda crise econômica. Ele e seu partido político decidiram que a estratégia eleitoral seria conquistar voto sustentando um discurso que culpa uma minoria pelos males da sociedade, principalmente pela crise econômica. A eficácia dessa estratégia eleitoral depende que os eleitores acreditem na contraposição proposta no discurso entre a minoria, apontada como causa dos problemas, e o agente político, apresentado como portador da solução. Supondo também que a contraposição seja uma representação falsa da realidade complexa dessa sociedade tanto pelo lado do diagnóstico do problema como da proposta de solução, reparemos que a base do discurso é a crença a respeito da causa do problema. Sem ela, a estratégia eleitoral não funcionaria. Ela seria, portanto, a crença de fundo a ser protegida pelo que vamos chamar de *estratégia desinformativa* do agente político.

A *estratégia desinformativa*, nesse caso, teria dois objetivos: 1) criar um conceito estereotipado – uma representação falsa, exagerada e caricata – na mente do eleitor a respeito da minoria de que trata seu discurso político, ou caso o conceito já esteja estabelecido e compartilhado, reforçá-lo; 2) criar uma crença – novamente, uma representação falsa – a respeito das causas dos problemas da sociedade, especificamente a crise econômica, ou, caso a crença já esteja estabelecida, reforçá-la e – ponto

essencial para nossa análise – protegê-la. Estamos diante de duas funções semiótico-comunicacionais introduzidas na seção 2, a função de *formação, uso e fixação de conceitos* e a função de *formação, uso e fixação de crenças* (respectivamente funções “d” e “e”). Por óbvio, as demais funções – *captura da atenção* (“a”), *manutenção da atenção/engajamento* (“b”), *orientação do processo interpretativo* (“c”) – terão que ser atendidas.

Tratemos do cenário em que tanto o conceito falso sobre a minoria como a crença a respeito da causa do problema econômico já estejam previamente estabelecidos nessa sociedade e sejam compartilhados entre muitos indivíduos. No caso de uma corrida eleitoral, o mais comum é que agentes políticos inescrupulosos recorram a preconceitos já arraigados, concepções estereotipadas, discursos de ódio e teorias da conspiração que já circulem na sociedade. Por critério de eficiência, parece que nesses casos estrategistas políticos optam por “não inventar a roda”.

O diagrama (figura 2) representa a dinâmica da circulação e interpretação de mensagens informativas e desinformativas no caso da estratégia eleitoral que estamos analisando. Os círculos, cheios ou vazados, representam as mensagens que podem circular nas duas esferas acima referidas, o retângulo pontilhado representa a esfera ampla, e o retângulo com linha cheia representa a esfera restrita (ambiente imediato). O último retângulo, que tem linha pontilhada dupla, representa o âmbito do processo interpretativo individual (este é o “ponto de chegada” das mensagens). As setas, pontilhadas ou com linha cheia, significam o caminho que a mensagem perfaz. Os “x” em vermelho representam o bloqueio; as pequenas barras vermelhas representam desvios no processo interpretativo (em relação ao curso normal, cf. seção 1). O processo interpretativo é representado pelo pequeno triângulo azul (juntamente com a seta que sai dele).

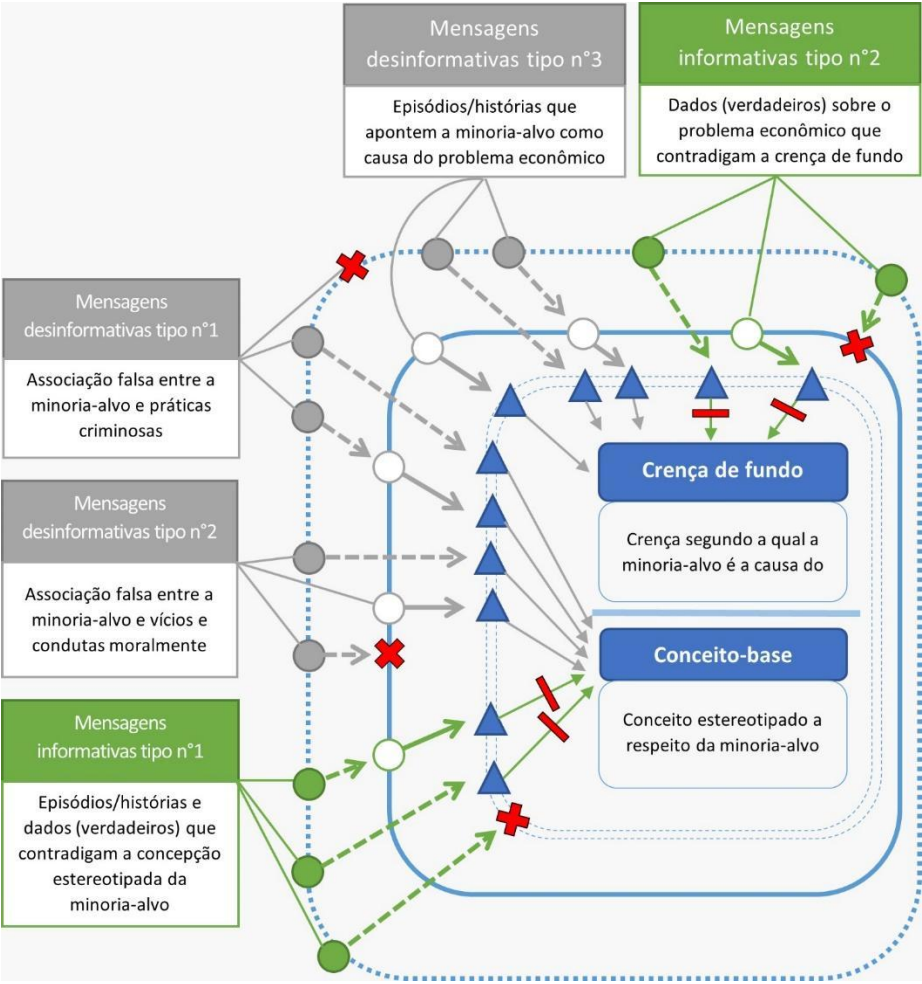


Figura 2 – Diagrama das dinâmicas de circulação de representações informativas e desinformativas (Fonte: elaborado pelo autor)

Os desvios (e eventualmente os bloqueios totais) em relação a uma atualização informativa dependem do grau de convicção do indivíduo (receptor/intérprete) na crença de fundo e do grau de

sedimentação do conceito-base. Tanto essa crença como o conceito são direcionados contra a minoria-alvo e constituem a causa imediata do desvio. Em termos semióticos, o desvio se realiza na alteração do peso da informação que chega pelos signos, isto é, na alteração da evidência relacionada à crença de fundo. A mensagem favorável tende a ter o peso aumentado para reforçar a crença de fundo (representada, no diagrama [figura 2], pela cor verde). Aquela que for desfavorável tende a ter o peso diminuído ou anulado com a mesma finalidade de reforçar a crença de fundo ou pelo menos diminuir o impacto (representada pela cor cinza).

Embora, do ponto de vista interpretativo, o “jogo” da desinformação seja, de fato, jogado no nível individual, os pontos de bloqueio de mensagens operam verdadeiros processos de filtragem também no nível social imediato e amplo (representados no diagrama pelos dois retângulos mais externos). Se observarmos o fenômeno da desinformação “de cima para baixo”, notaremos que essa filtragem na comunicação de larga escala estabelece condições para uma eficiente forma vertical de reforço de uma crença socialmente compartilhada, o que, por sua vez, retroalimenta uma forma horizontal de reforço (dessa mesma crença), que é aquela que opera no nível social imediato (grupos primários, como famílias, ambiente de trabalho, instituições comunitárias etc.). Observando “de baixo para cima”, a filtragem na comunicação do ambiente social imediato cria as condições para que os efeitos acumulados dessa forma horizontal de reforço retroalimentem de forma eficiente a forma vertical acima descrita. Sem contraposição, esses *loopings* (essas alças de retroalimentação) estão destinados a gerar um grau de radicalização de crenças e sedimentação de conceitos capaz de inviabilizar nossas democracias liberais. Quando as concebemos éramos mais ingênuos tanto em relação às nossas capacidades cognitivas quanto aos efeitos de nossas parafernalias tecnológicas.

Reparemos que muitos dos truques mobilizados pelas estratégias desinformativas típicas do cenário contemporâneo da comunicação são velhos conhecidos, alguns deles foram capturados e sistematizados ainda no século IV a.C. por Aristóteles na obra *Retórica*, muitos deles sempre foram utilizados por demagogos e charlatães para enganar o público incauto independente da época, do espaço e do meio utilizado. A novidade é que fenômeno da desinformação – conforme descrito nestas páginas – é o resultado de séculos de experiência acumulada de uma *retórica da covardia* desenvolvida para explorar fragilidades cognitivas humanas e que, nesses últimos anos, tem sido potencializada por tecnologia de ponta atingindo escala industrial. Assim, defendemos que desinformação é retórica da covardia anabolizada por tecnologia comunicacional.

Referências

- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139.
- Bond, R.; Fariss, C.; Jones, J.; Kramer, A.; Marlow, C.; Settle, J. & Fowler, J. (2012) A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415).
- Desmurget, M. (2021). *A fábrica de cretinos digitais: os perigos das telas para nossas crianças*. Vestígio.
- Fallis, D. (2015). What Is Disinformation? *Library Trends*, 63(3), 401–426.
- Gillespie, T. (2014). The Relevance of Algorithms. *Light* 1999.
- Haack, S. (1982). Descartes, Peirce and the Cognitive Community. *The Monist*, v. 65, n. 2, p. 166-71.

- Haidt, J. (2024). *A geração ansiosa*: Como a infância hiperconectada está causando uma epidemia de transtornos mentais. Companhia das Letras.
- Kaiser, B. (2019). *Targeted*: My inside story of Cambridge Analytica and how Trump and Facebook broke democracy. Harper Collins.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697-720.
- Kahneman, D. (2011). *Thinking fast and slow*. Penguin Books.
- Kahneman, D; Sibony, O. & Sunstein, C. (2021). *Ruído*: uma falha no julgamento humano. Objetiva.
- Kahneman, D. & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-291.
- Matthews, R. (2017). *As leis do acaso*: como a probabilidade pode nos ajudar a compreender a incerteza. Zahar.
- Matz, S. C.; Kosinski, M.; Nave, G. & Stillwell, D. J. (2017) Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the national academy of sciences*, 114(48).
- Oliveira, T. (2020). Desinformação científica em tempos de crise epistêmica: Circulação de teorias da conspiração nas plataformas de mídias sociais. *Fronteiras - estudos midiáticos*, 22(1).
- Pariser, E. (2012). *O filtro invisível*: O que a internet está escondendo de você. Zahar.
- Peirce, C. (1877). The fixation of belief. In: Peirce, C. S. (1931). *The Collected Papers of Charles Sanders Peirce: Vol. I-VIII (1931-1978)* (C. Hartshorne, P. Weiss, & A. Burks, Orgs.) [CP 4.358-387].

- Recuero, R. (2020). *SciELO - Brasil—#FraudenasUrnas*: Estratégias discursivas de desinformação no Twitter nas eleições 2018. <https://www.scielo.br/j/rbla/a/vKngHPRMJxbypBVRLYN3YTB/?lang=pt>
- Santaella, L. (2023). *Definir desinformação é preciso*. In: Santaella, L. Flagelos da desinformação (p. 25–54). Educ.
- Shannon, C. E. & Weaver, W. (1975) *A teoria matemática da comunicação*. Difel.
- Simon, H. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1), 99-118.
- Simon, H. (1982). *Models of bounded rationality*. Cambridge, MA, MIT Press.
- Sunstein, C. (2020). *Behavioral Science and Public Policy*. Cambridge Books, Cambridge University Press.
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, Elsevier, 1(1), pages 39-60.
- Thaler, R. & Sunstein, C. (2021). *Nudge: The Final Edition*. Penguin Press.
- Wardle, C. & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe.