

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE FILOSOFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

Rogério Fernandes Martins

A razão comprometida: grupos polarizados e explicações radicais

Versão Corrigida

São Paulo
2023

Rogério Fernandes Martins

A razão comprometida: grupos polarizados e explicações radicais

Versão Corrigida

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia do Departamento de Filosofia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Mestre em Filosofia sob a orientação do Prof. Dr. Caetano Ernesto Plastino

São Paulo
2023

ENTREGA DO EXEMPLAR CORRIGIDO DA DISSERTAÇÃO/TESE

Termo de Anuência do (a) orientador (a)

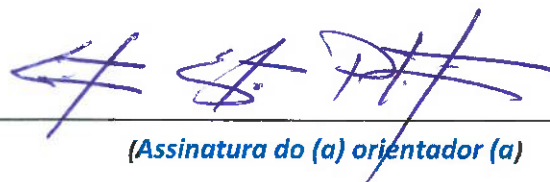
Nome do (a) aluno (a): Rogério Fernandes Martins

Data da defesa: 17/10/2023

Nome do Prof. (a) orientador (a): Caetano Ernesto Plastino

Nos termos da legislação vigente, declaro **ESTAR CIENTE** do conteúdo deste **EXEMPLAR CORRIGIDO** elaborado em atenção às sugestões dos membros da comissão Julgadora na sessão de defesa do trabalho, manifestando-me **plenamente favorável** ao seu encaminhamento ao Sistema Janus e publicação no **Portal Digital de Teses da USP**.

São Paulo, 04/12/2023



(Assinatura do (a) orientador (a))

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

M379r Martins, Rogério Fernandes
 A razão comprometida: grupos polarizados e
 explicações radicais / Rogério Fernandes Martins;
 orientador Caetano Ernesto Plastino - São Paulo,
 2023.
 142 f.

Dissertação (Mestrado)- Faculdade de Filosofia,
Letras e Ciências Humanas da Universidade de São
Paulo. Departamento de Filosofia. Área de
concentração: Filosofia.

1. Polarização grupal. 2. Explicações. 3. Bolhas
epistêmicas. 4. Formação de crenças. 5. Epistemologia
social. I. Plastino, Caetano Ernesto, orient. II.
Título.

Folha de Aprovação

MARTINS, R. F. **A razão comprometida**: grupos polarizados e explicações radicais. 2023. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo, 2023.

Aprovado em: _____

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Julgamento _____

Prof. Dr. _____

Instituição: _____

Julgamento _____

Prof. Dr. _____

Instituição: _____

Julgamento _____

Prof. Dr. _____

Instituição: _____

Julgamento _____

Dedico este trabalho à memória do meu pai, Antônio. Que saudade de você, meu jovem! Muito da minha dedicação à sua elaboração foi motivada pela tentativa de superar a sua falta.

Agradecimentos

À minha amada esposa, Tânia, agradeço pela paciência, amor, compreensão e pelo incentivo constante. Sem o seu apoio, esta conquista não seria possível. Obrigado de todo o meu coração!

À minha mãe, Marlene, cuja dedicação e amor são um exemplo para mim. Você me ensinou tanto! Principalmente, a persistir e não desistir diante dos desafios. Obrigado por acreditar em mim quando eu mais precisava!

Ao meu filho, Guilherme. A sua presença na minha vida é uma constante fonte de alegria e orgulho. Além de ser um filho muito dedicado (às vezes, um pouco teimoso!), você é, acima de tudo, meu amigo!

À minha nora, Vanessa. Agradeço pelas palavras de incentivo e carinho que sempre me ofereceu quando eu expressava minhas dificuldades e dúvidas. Acredite, elas foram muito importantes!

Finalmente, ao Professor Caetano, um mestre excepcional, tão querido e admirado por todos. Que honra e que privilégio foi tê-lo como orientador!

Verdade e mentira têm igual fisionomia; vemo-las com os mesmos olhos. Não somente nos mostramos covardes e nos defendemos com tão pouco ardor contra a impostura, como ainda nela chafurdamos; apraz-nos emaranhar-nos em vaidade, como se ela fizesse parte de nosso ser.

(MONTAIGNE, 2016, p.947)

RESUMO

MARTINS, R. F. **A razão comprometida**: grupos polarizados e explicações radicais. 2023. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo, 2023.

Esta dissertação de mestrado examina, sob uma perspectiva filosófica, aspectos da intensa polarização de crenças e opiniões presentemente observada em diversos grupos sociais. Mais precisamente, a análise se concentra na persistência de discursos explicativos falaciosos dentro desses grupos, expandindo-se em um estudo abrangente sobre sua origem e perpetuação. Inicialmente, o termo "polarização" é definido como o deslocamento de opinião para extremos em grupos compostos por indivíduos que pensam de modo semelhante, após a ocorrência de algum evento deliberativo. Particularmente, a atenção concentra-se nos grupos polarizados mais coesos e persistentes, denominados "bolhas epistêmicas". Estes são ambientes nos quais os indivíduos não conseguem distinguir entre o que acreditam conhecer e o que de fato conhecem. Nomeia-se o discurso explicativo falacioso proferido nesses contextos como "razão comprometida". Essa denominação abrange duas interpretações distintas. A primeira interpretação considera a expressão "razão comprometida" como um raciocínio duvidoso ou falho, levantando a questão: por que explicações ilegítimas são tão persuasivas no interior dessas bolhas? Para examinar essa situação, recorre-se a uma análise combinada de várias teorias, incluindo a teoria pragmática da explicação de van Fraassen, a teoria dos atos de fala de Austin e a teoria das implicaturas de Grice. Essa análise revela que a força e os efeitos produzidos por uma explicação não dependem necessariamente da veracidade de seus componentes. A segunda interpretação considera a expressão "razão comprometida" como o resultado de um compromisso grupal, explicando a sustentação e a tenacidade na defesa de discursos falaciosos, mesmo que fortemente contestados, no interior desses grupos. Nessa abordagem, recorre-se à epistemologia dos grupos e à tese da negociação de crença coletiva de Gilbert e Priest, que afirma que as crenças grupais emergem de uma negociação linguística coletiva, culminando em um compromisso mútuo com o que é proferido nesse contexto. Isso sugere a prevalência de lealdade grupal sobre razões epistêmicas na articulação, por exemplo, de discursos explicativos no interior desses grupos. As considerações desenvolvidas ao longo do texto facilitam a análise e a avaliação das intervenções atualmente propostas para mitigar os efeitos da polarização, além de sugerirem um caminho para futuras pesquisas.

Palavras-chave: Polarização grupal. Explicações. Bolhas epistêmicas. Formação de crenças. Epistemologia social.

ABSTRACT

MARTINS, R. F. **Compromised reason**: polarized groups and radical explanations. 2023. Thesis (Master Degree) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo, 2023.

This master's thesis examines, from a philosophical perspective, aspects of the intense polarization of beliefs and opinions presently observed in diverse social groups. More precisely, the analysis focuses on the persistence of fallacious explanatory discourses within these groups, expanding into a comprehensive study of their origin and perpetuation. Initially, the term "polarization" is defined as the shift in opinion towards extremes in groups composed of like-minded individuals following the occurrence of a deliberative event. Particular attention is paid to the most cohesive and persistently polarized groups, termed "epistemic bubbles". These are environments in which individuals are unable to distinguish between what they believe they know and what they actually know. The fallacious explanatory discourse presented in these contexts is referred to as "compromised reason". This term encompasses two distinct interpretations. The first interpretation considers the expression "compromised reason" as dubious or flawed reasoning, raising the question: why are illegitimate explanations so persuasive within these bubbles? To examine this situation, an analysis that combines several theories is employed, including van Fraassen's pragmatic theory of explanation, Austin's speech act theory, and Grice's theory of implicatures. This analysis reveals that the strength and effects produced by an explanation do not necessarily depend on the truthfulness of its components. The second interpretation considers the expression "compromised reason" as the result of a group commitment, explaining the support and tenacity in defending fallacious discourses, even when they are heavily contested within these groups. In this approach, the study refers to the epistemology of groups and Gilbert and Priest's thesis on the negotiation of collective belief, asserting that group beliefs emerge from a collective linguistic negotiation, culminating in a mutual commitment to the utterances made within this context. This suggests the prevalence of group loyalty over epistemic reasons in the articulation of explanatory discourses within these groups, for instance. The considerations developed throughout the text facilitate the analysis and evaluation of interventions currently proposed to mitigate the effects of polarization while also suggesting a path for future research.

Keywords: Group polarization. Explanations. Epistemic bubbles. Belief formation. Social epistemology.

SUMÁRIO

1	INTRODUÇÃO	10
2	O FENÔMENO DA POLARIZAÇÃO GRUPAL	15
2.1	A lógica da polarização grupal	27
2.2	As causas da polarização grupal	28
2.3	A polarização grupal e as bolhas epistêmicas	30
3	EXEMPLOS DE BOLHAS EPISTÊMICAS	33
3.1	A doutrina da bruxaria no período renascentista	33
3.2	Winston Parva	42
3.2.1	A fofoca: sua função social e epistêmica	45
3.3	As mídias sociais	50
4	OS MODELOS EXPLICATIVOS	61
4.1	O modelo nomológico-dedutivo	63
4.2	O modelo de relevância estatística	67
4.3	Os aspectos pragmáticos envolvidos numa explicação	70
5	A TEORIA PRAGMÁTICA DA EXPLICAÇÃO DE van FRAASSEN	74
5.1	Um exemplo de uma explicação pragmática	83
5.2	Os pressupostos de uma questão-por-quê	85
6	A TEORIA DOS ATOS DE FALA E A TEORIA DAS IMPLICATURAS	90
6.1	Como um discurso explicativo comprometido pode ser tão efetivo	104
7	BOLHAS E DESACORDOS EPISTÊMICOS	109
8	A EPISTEMOLOGIA DOS GRUPOS	119
8.1	A tese da negociação de crença coletiva	127
9	CONCLUSÃO	131
	REFERÊNCIAS	137

1 INTRODUÇÃO

Interpretar o significado do pequeno conjunto de palavras que compõem o título e o subtítulo deste trabalho pode ser um bom caminho para introduzir o tema que trataremos ao longo do texto. Este simples exercício interpretativo não demandaria muito esforço nem grande engenho, especialmente para aqueles que possuem o privilégio¹ de poder se dedicar ao exame e à reflexão das nuances de eventos que, embora imperceptíveis para a maioria das pessoas, acabam por moldar a nossa realidade.

A princípio, não seria muito difícil inferir que a presente dissertação trata de um fenômeno que tem atraído a atenção e preocupado muitas pessoas recentemente: a polarização das crenças e das opiniões. Nesse contexto, tendo em vista aquele breve exercício de interpretação, acreditamos que termos como "*Fake News*", "*WhatsApp*", "manipulação", "discursos de ódio", "grupos polarizados" e "radicalização" teriam uma boa chance de surgir na mente de um leitor mais atento e preocupado. Mais do que meras palavras e expressões, particularmente para este leitor angustiado, é bastante razoável supor que esses termos venham a despertar em seu ser uma grande variedade de sentimentos. Entre estes, talvez o receio de que as atuais mídias sociais possam ser as principais responsáveis pelo incerto momento que presenciamos, atuando como verdadeiras máquinas polarizadoras, sob o comando de poderosos algoritmos oniscientes. Contudo, haveria razão para tal preocupação?

Para começar, conforme entendemos, e a interpretação do título sugere, a polarização grupal é um fato. Ela está fragmentando a opinião pública, ao extremar discursos, deteriorando assim a razão. Entretanto, para sermos precisos, o fenômeno da polarização grupal não é um sintoma do nosso tempo. Trata-se de um evento comum, e a radicalização que algumas vezes o acompanha tem sido frequentemente testemunhada ao longo da história. No entanto, proporcionado pelo desenvolvimento das atuais mídias sociais, o grande número de grupos polarizados, muitas vezes radicalizados, e, por essa razão, o retorno de temas ultrapassados do ponto de vista civilizatório, são fatos inéditos e muito preocupantes.

¹ Notadamente, em um país onde uma parte considerável de seu povo é tão desassistida em aspectos tão básicos quanto a alimentação, por exemplo.

Sabe-se que, desde o final do século XIX, os fenômenos de massa (e a polarização grupal encontra-se entre eles) têm sido objeto de estudo de psicólogos, psicanalistas e sociólogos. Contudo, apenas muito recentemente esses fenômenos passaram a despertar a atenção dos filósofos. Inclusive, até mesmo aqueles cujas pesquisas não se concentram nesse tema, já perceberam que algumas concepções até então bem consolidadas em seus campos de estudo, diante da atual difusão de grupos polarizados, necessitam no mínimo de reformulação. Ecos dessa preocupação são o que podemos constatar nas palavras da professora e filósofa Marilena Chaui (2022) em palestra recente:

[Exibidas como um novo traço da ideologia contemporânea, as bolhas representam] uma nova forma de sociabilidade, porque na ideologia tradicional você se referia a sociedade como um todo, o que era [um fato] fundamental, indivisa, una, centrada etc. [...]. Agora não, ela está toda dividida, mas não em classes, mas dividida em bolhas eletrônicas que produzem a ilusão da sociabilidade e a verdade da solidão e da competição. Uma coisa fora do comum! (informação verbal)²

O trabalho que ora apresentamos é um pequeno passo no sentido de um tratamento filosófico do tema. A partir de um espectro de teorias que vão desde formulações na área da filosofia da ciência, passando pela filosofia da linguagem e culminando em estudos recentes em epistemologia social, objetivamos apreender aspectos desse estado de coisas que, embora não sendo originais, se intensificaram de forma bastante preocupante, como mencionamos.

Numa primeira aproximação, podemos pensar que, quando um grupo se polariza, o discurso racional compartilhado por seus integrantes acompanha tal movimento, enfraquecendo e até falseando os seus modos clássicos de apresentação: as asserções, os argumentos e as explicações. Ou seja, conforme explicitamos no título do trabalho, **a razão se compromete no interior daqueles grupos**. No sentido de se compreender esse fato, concentraremos especialmente a nossa atenção no estudo do discurso explicativo falacioso. *Grosso modo*, enquanto um argumento tem como objetivo demonstrar e justificar a verdade de uma determinada proposição, uma explicação é um raciocínio que visa apresentar causas

² CHAUI, Marilena. **O que é ideologia? Palestra com Marilena Chaui e Jessé Souza**. YouTube, 08 de abr. de 2022. Disponível em: https://www.youtube.com/watch?v=C3wv_vpRjzk. Acesso em: 18 de dez. de 2022.

ou razões para a ocorrência de um fato que sabemos de antemão ser o caso. Por que Portugal foi o pioneiro na grande expansão marítimo-comercial que se iniciou no século XV? Por que a água congelou a zero graus Celsius naquele recipiente? Por que o Brasil foi o último país a abolir a escravidão? Por que Wittgenstein abandonou o atomismo lógico no final da década de 1920? Nota-se como, ao procurarmos responder a essas questões, buscamos no fundo elucidar, ou melhor, explicar eventos que indubitavelmente ocorreram.

Ademais, o discurso explicativo apresenta um aspecto ainda mais fundamental: ele revela como as pessoas, ao atribuir sentido aos fatos, percebem e interpretam os acontecimentos do seu entorno. Especificamente, considerando os objetivos deste estudo, as explicações oriundas de um grupo social refletem boa parte da base coletiva de crenças que fundamenta o modo como esses indivíduos compreendem e atuam sobre a realidade, conforme veremos.

Adicionalmente, ao se refletir sobre os efeitos de uma explicação, vale notar a existência de outro fato essencial: “explicar algo é um ato epistêmico; quando algo lhe é explicado, você adquire conhecimento” (LYCAN, 2002, p.408, tradução nossa). Diante desse vínculo entre explicar e conhecer, esperava-se, por exemplo, que explicações que desrespeitassem a normatividade epistemológica tradicional, provavelmente não se sustentassem. Contudo, os discursos explicativos falaciosos são muito influentes principalmente no interior de grupos polarizados. Ou seja, a **“razão comprometida” é muito persuasiva**. Aspectos relacionados às condições de verdade desses discursos parecem não fazer muita diferença, importando mais a sua aparência de correção e o seu poder de persuasão.

Por fim, outro importante aspecto a ser observado é que mesmo que frequentemente desafiados, tais discursos são sustentados e defendidos com muita tenacidade por aqueles grupos. Seus membros parecem acreditar firmemente nessas explicações. O que nos direciona, em uma nova chave interpretativa, para o outro sentido do adjetivo “comprometida”. **A expressão “razão comprometida” passa a ser agora compreendida enquanto figuração e expressão de um compromisso entre agentes**. Compromisso em acatar, respeitar e defender crenças que, uma vez acordadas, sustentariam e caracterizariam respectivamente a coesão e a identidade grupal. A partir dessa nova compreensão, mostraremos que, contrapondo-se às possíveis considerações de caráter epistêmico, existiria uma motivação de lealdade

às crenças acordadas no interior desses grupos polarizados. Isso explicaria, entre outras coisas, a sustentação e a defesa de posicionamentos grupais frequentemente desmedidos e irracionais.

Vale observar que à medida que procuramos compreender o fenômeno da polarização, delimitamos também as possíveis causas desse estado de coisas. Dessa forma, cremos que as formulações discutidas neste estudo, bem como suas descobertas, poderiam auxiliar na avaliação de medidas destinadas a combater a radicalização e seus efeitos.

Em breve apresentação, estes são os capítulos que se dedicarão a enfrentar a complexa missão a que nos propomos. O primeiro capítulo, intitulado "**O Fenômeno da Polarização Grupal**", fornecerá uma visão geral sobre esse fenômeno. Estudaremos a polarização sob a perspectiva dos estudos mais recentes nos campos da psicologia cognitiva e social. A apresentação do importante conceito de bolha epistêmica enquanto um grupo polarizado persistente concluirá o capítulo.

O segundo capítulo, "**Exemplos de Bolhas Epistêmicas**", ilustrará o conceito de bolha epistêmica por meio da análise de três casos: a caça às bruxas na Renascença; a discriminação e estigmatização social ocorridas em uma pequena localidade na Inglaterra entre as décadas de 1950 e 1960; e a perturbadora situação atual das mídias sociais.

Considerando que um de nossos objetivos é a análise do comportamento do discurso explicativo, o terceiro capítulo, "**Os Modelos Explicativos**", apresenta as principais propostas surgidas sobre esse tema na pesquisa filosófica desde meados do século XX. Conforme planejamos mostrar, os modelos mais adequados aos nossos propósitos argumentativos são aqueles que enfatizam, entre outros elementos, a importância das circunstâncias de fala durante a ocorrência de uma interação explicativa. Por esse motivo, optamos por utilizar o modelo desenvolvido pelo filósofo Bas C. van Fraassen, que será detalhado no capítulo "**A Teoria Pragmática da Explicação de van Fraassen**".

No capítulo "**A Teoria dos Atos de Fala e a Teoria das Implicaturas**", pretendemos alcançar o nosso primeiro grande objetivo analítico e responder à questão: Por que a razão, mesmo configurada como um discurso explicativo notadamente comprometido, pode ainda ser tão persuasiva e eficaz?

Os dois capítulos finais, "**Bolhas e Desacordos Epistêmicos**" e "**A Epistemologia dos Grupos**", em que estudaremos recentes pesquisas na área da epistemologia social, concluem nosso estudo. Neles, nossa atenção se concentrará especialmente no segundo sentido da expressão "razão comprometida", conforme assinalamos acima.

Finalmente, na "**Conclusão**", além da apresentação dos destaques e descobertas do estudo, discutiremos brevemente algumas pesquisas atuais sobre o importante tema da despolarização grupal.

2 O FENÔMENO DA POLARIZAÇÃO GRUPAL

Integramos vários grupos durante a nossa vida, e esse fato suscita consequências. Por exemplo, ser servidor público sindicalizado; ou fazer parte da comunidade de moradores de Heliópolis; ou mesmo seguir em uma rede social as discussões do Instituto de Economia da UFRJ; implicam necessariamente determinados níveis de comprometimento que vão desde a participação em deliberações formais até a defesa de certos posicionamentos.³

Uma vez que um razoável procedimento de avaliação e de escolha esteja sendo cumprido, respeitando principalmente à livre manifestação de opinião, não seria incomum aqueles agentes decidirem em uma assembleia trabalhista pela paralisação de um dia de trabalho; aprovarem um plano de construção para uma nova quadra de esportes em uma reunião de moradores; ou ainda endossarem, em alguma mídia social, uma crítica à atual política de preços de combustíveis da Petrobras. No entanto, muitas vezes, desmedidamente, aprova-se uma greve por tempo indeterminado; decide-se pela construção de dez novas quadras; inicia-se uma campanha de ódio contra o atual presidente da Petrobras nas mídias sociais. Em tais situações, dizemos habitualmente que o grupo se polarizou ou, em casos extremos, que se radicalizou.

Valendo-se de uma terminologia apresentada por Tsebelis (1998)⁴, a surpresa dessas ações estaria no fato de que elas divergem de um certo comportamento racional que nós, observadores, indivíduos não pertencentes a esses grupos, consideraríamos adequado ou ótimo. Ou seja, enquanto observadores, podemos achar a atuação de determinados agentes completamente irracional. Entretanto, tais atores acreditam que suas ações seguem um plano racional.⁵

³ Vale notar que um determinado indivíduo poderia pertencer simultaneamente a todos esses grupos, por exemplo.

⁴ Conforme Tsebelis, o ponto de vista é crucial na avaliação de uma ação racional. Como observadores, podemos considerar a atuação de determinados atores ou agentes como irracional. Entretanto, esses atores podem estar seguindo um plano racional que muitas vezes se mostra opaco às nossas observações. O autor utiliza essa ideia na sua obra "Jogos Ocultos: Escolha racional no campo da política", na qual procura analisar o comportamento aparentemente irracional de agentes e partidos políticos.

⁵ Karl Popper (1996, 2004) sugere a "análise situacional" ou a 'lógica situacional' como um método para entender, em contextos complexos, a racionalidade das ações de indivíduos que, a princípio, pareceriam caóticas e desprovidas de qualquer planejamento aparente. O filósofo argumenta que as ações são racionais quando são adequadamente ajustadas à situação. Na obra "O mito do contexto", Popper (1996) ilustra o método, ao analisar o comportamento aparentemente errático de um pedestre

Conforme vimos nos exemplos apresentados no início deste capítulo, vale observar como a polarização grupal necessita de um evento deliberativo presencial ou mesmo virtual para existir. A princípio, a ideia de que uma deliberação possa gerar uma polarização grupal não deixa de causar um certo estranhamento. Ora, de acordo com Simon (1990, p.7, tradução nossa), uma vez que o comportamento racional “[...] molda-se por uma tesoura cujas lâminas são respectivamente a estrutura ambiental [contextual], onde se desenrola uma determinada tarefa, e as capacidades computacionais [cognitivas] do agente”; seria razoável supor que a execução de um processo decisório em grupo, por si só, já garantisse uma melhor avaliação e um melhor resultado. Afinal, em uma deliberação grupal, não teríamos mais olhos para examinar todas as nuances da informação fornecida pelo ambiente bem como uma maior capacidade cognitiva disponível para um julgamento mais preciso? É notório o paralelismo entre o conteúdo dessa questão e um certo modelo de racionalidade instrumental bastante disputado. O modelo: o *homo economicus*. Segundo Edwards (1954), um agente racional nesse modelo:

- i. encontrar-se-ia completamente informado sobre os aspectos relacionados ao assunto em questão. Isto é, todas as opções disponíveis, bem como todas as consequências de cada decisão, seriam passíveis de análise;
- ii. seu modo de agir seria infinitamente sensível às nuances dessas opções e dessas consequências;
- iii. por fim, tal agente conseguiria categorizar, quantificar e ordenar as alternativas disponíveis, escolhendo sempre o melhor curso da ação.

Nesse sentido, se imaginarmos a deliberação no interior de um grupo infinito de indivíduos, tal processo decisório levado ao limite por essas circunstâncias, não aproximaria as decisões grupais do comportamento ideal do *homo economicus*? Ademais, ao se projetar tal idealização no plano da realidade, justificar-se-ia a

que atravessa uma rua movimentada. Sem levar em conta as motivações psicológicas do indivíduo, a análise situacional considera objetivamente tanto os elementos físicos e sociais presentes no cenário da ação quanto o conhecimento que o agente tem deles. Embora forneça um modelo racional passível de verificação empírica, a análise situacional não reivindica precisão absoluta. Ela seria uma aproximação útil; todavia, não exaustiva da realidade. Para os propósitos do nosso estudo, ao não considerar os aspectos psicológicos dos agentes, esse método apresenta limitações. Como veremos adiante, em cenários de acentuada polarização, fatores subjetivos exercem papéis significativos no comportamento coletivo.

seguinte crença comum: quanto mais integrantes em deliberação, maior a probabilidade da certeza dos atos e menores os riscos da ocorrência de um julgamento indevido. Aristóteles (1985), na sua obra, “Política”, já afirmara esse fato:

É realmente possível que os integrantes da maioria, apesar de não serem isoladamente homens bons, possam ser melhores quando reunidos, não individualmente, mas coletivamente, do que os poucos individualmente bons, da mesma forma que os banquetes públicos, para os quais muitos contribuem, são melhores que os oferecidos a expensas de um só homem; realmente pode-se argumentar que onde há muitos homens, cada um tem uma certa porção de bondade e prudência, e que quando eles se reúnem, da mesma maneira que a multidão se transforma num homem só com muitos pés e muitas mãos e muitos sentidos, podem assumir também uma só personalidade quanto às faculdades morais e intelectuais. É esta a razão de a maioria ser melhor juiz para obras musicais e a dos poetas, pois homens diferentes podem julgar aspectos diferentes do desempenho e todos eles podem julgar o conjunto. (ARISTÓTELES, 1985, p.98)

O teorema do júri, proposto pelo Marquês de Condorcet em seu trabalho de 1785, “Ensaio sobre a aplicação da análise de probabilidade às decisões tomadas por pluralidade de votos”, demonstra matematicamente a mesma concepção aristotélica: quanto maior o número de indivíduos em deliberação, maior a probabilidade da ocorrência de um julgamento justo.⁶ Entretanto, como podemos explicar que, à

⁶ Para uma compreensão prática do teorema, examinemos o seguinte exemplo em que analisamos a ocorrência de um julgamento em um tribunal do júri em primeira instância. Para começar, supomos que temos as seguintes informações:

- Habitualmente, um mesmo conjunto de jurados participa dos julgamentos naquele tribunal;
- Cada jurado que integra esse conjunto avalia as evidências disponíveis de forma independente e com adequada competência. Quando analisamos o histórico de cada um deles, independentemente do réu ser inocente ou culpado, sabemos que eles acertam 60% dos vereditos, errando, assim, 40% deles. Esses valores foram calculados com base na promulgação das sentenças finais em trânsito em julgado dos casos em que aqueles jurados participaram em primeira instância. Toda vez que o conteúdo (culpado ou inocente) de um veredito de um jurado em primeira instância é confirmado na sentença final em trânsito em julgado, computa-se um acerto em seu histórico de vereditos; caso contrário, um erro;
- Apenas 3 jurados daquele conjunto participam do julgamento.

Apresentadas essas informações, qual seria a probabilidade desse julgamento apresentar um resultado justo, independentemente da inocência ou da culpa do réu?

Para que haja um resultado justo no mínimo dois dos três jurados devem acertar o resultado do veredito. Tendo isso em vista, percebemos que existem somente quatro possibilidades que atendem a essa condição, a saber:

- 1) o jurado 1 acerta, o jurado 2 acerta e o jurado 3 erra.
- 2) o jurado 1 acerta, o jurado 2 erra e o jurado 3 acerta.
- 3) o jurado 1 erra, o jurado 2 acerta e o jurado 3 acerta.
- 4) todos os jurados acertam os seus vereditos individuais.

margem de qualquer parâmetro racional ou até civilizatório, alguns grupos, mesmo contando com a participação de muitos agentes em deliberação, possam se polarizar de maneira tão extrema? Por que a maquinaria racional tradicional, que em tese deveria estar em ação nessas conjunturas, se mostra tão inoperante?

Uma primeira tentativa de explicação concentra-se no estudo das deficiências dos atores grupais e nas suas possíveis ações. Mais especificamente, o campo de estudos da psicologia cognitiva procura explicar, entre outras coisas, essas supostas deficiências. Consoante um primeiro conjunto de teorias, as possíveis falhas nos processos de deliberação grupal poderiam ser explicadas a partir de problemas de coordenação entre as intuições e as capacidades cognitivas de ordem superior dos agentes. De acordo com os autores dessa visão, tais problemas de coordenação originam os chamados “vieses cognitivos”. Estes representam erros de julgamento que ocorrem principalmente quando nossas avaliações se baseiam em heurísticas. As heurísticas são atalhos mentais, regras simplificadas de raciocínio que nos auxiliam na resolução de problemas quando não dispomos de informações suficientes. O exemplo a seguir, mostra-nos a possível ocorrência de uma falha de raciocínio, ou melhor, de um viés cognitivo, quando executamos a chamada heurística da representatividade:

[...] “Steve é muito tímido e retraído, invariavelmente prestativo, mas com pouco interesse nas pessoas ou no mundo real. De índole dócil e organizada, tem necessidade de ordem e estrutura, e uma paixão pelo detalhe.” Como as pessoas avaliam a probabilidade de que Steve esteja envolvido em uma ocupação particular de uma lista de possibilidades (por exemplo, fazendeiro, vendedor, piloto comercial, bibliotecário ou médico)? Como as pessoas ordenam essas

E assim, levando-se em consideração as probabilidades dos acertos e dos erros de cada jurado, a probabilidade de um julgamento justo é:

$$P(\text{julgamento justo}) = (0,6 \times 0,6 \times 0,4) + (0,6 \times 0,4 \times 0,6) + (0,4 \times 0,6 \times 0,6) + (0,6 \times 0,6 \times 0,6) = 0,648 = 64,8\%$$

Vale notar que a probabilidade de 3 jurados acertarem o veredito final (64,8%) é maior que a de um só jurado acertar (60%). E se fossem 5 jurados? Nesse caso, realizando-se os cálculos, a probabilidade de 5 jurados acertarem o veredito (68,3%) seria ainda maior que a de um só jurado acertar (60%). Em um caso limite, com o número de jurados tendendo ao infinito, a probabilidade de um julgamento justo tenderia a 100%. O Teorema do Júri de Condorcet afirma que, quanto maior o número de jurados, maior a probabilidade de se obter um julgamento justo. Entretanto, a aplicação do teorema também indica que se a probabilidade de cada jurado acertar fosse de 50%, o aumento do número de jurados não impactaria a qualidade do julgamento. E, se a probabilidade de vereditos corretos de cada jurado fosse inferior a 50%, aumentar o número de jurados poderia piorar a qualidade do julgamento.

ocupações da mais para a menos provável? Na heurística da representatividade, a probabilidade de que Steve seja um bibliotecário, por exemplo, é avaliada segundo o grau em que ele é representativo, ou similar ao estereótipo de um bibliotecário. De fato, a pesquisa com problemas desse tipo mostrou que as pessoas ordenam as ocupações pela probabilidade e pela similaridade exatamente da mesma forma. Essa abordagem do julgamento da probabilidade leva a graves erros, pois a similaridade, ou representatividade, não é influenciada por diversos fatores que decerto afetarão os julgamentos de probabilidade. (KAHNEMAN, 2012, p.525)

É razoável supor que existam mais vendedores que bibliotecários no mundo. Entretanto, devido à heurística da representatividade, ignoramos esse fato, ao ancorar, por exemplo, nossas convicções de que Steve seja um bibliotecário na simples descrição de alguns traços de sua personalidade.

A visão mais conhecida e disputada nessa abordagem cognitiva é a teoria do processo dual de Kahneman. Segundo Kahneman (2003, 2012), em muitas circunstâncias do nosso dia a dia, realizamos julgamentos rápidos em ações que notadamente envolvem pouco esforço cognitivo. No entanto, também somos capazes de elaborar julgamentos mais elaborados, mais lentos, envolvendo significativo esforço nessa ação. A partir desses fatos, o autor desenvolveu a concepção de que lançamos mão de dois sistemas de processos cognitivos distintos em nossas atividades intelectuais, a saber: o Sistema 1 e o Sistema 2.

As operações do Sistema 1 “são rápidas, automáticas, sem esforço, associativas, implícitas (não disponíveis à introspecção) e frequentemente carregadas de emoção; também são governadas pelo hábito; portanto, difíceis de controlar ou modificar” (Kahneman, 2003, p.698, tradução nossa). Em ordem de complexidade, eis alguns exemplos das atividades realizadas pelo Sistema 1:

- Detectar que um objeto está mais distante que outro.
- Orientar-se em relação à fonte de um som repentino.
- Completar a expressão “pão com...”
- Fazer “cara de aversão” ao ver uma foto horrível.
- Detectar hostilidade em uma voz.
- Responder $2 + 2 = ?$
- Ler palavras em grandes cartazes.
- Dirigir um carro por uma rua vazia.
- Encontrar um movimento decisivo no xadrez (se você for um mestre enxadrista).
- Compreender sentenças simples.

- Reconhecer que uma “índole dócil e organizada com paixão pelo detalhe” se assemelha a um estereótipo ocupacional. (KAHNEMAN, 2012, p.30)

Por outro lado, as operações do Sistema 2 são “mais lentas, sequenciais, exigem esforço, mais propensas a serem conscientemente monitoradas e deliberadamente controladas; também são relativamente flexíveis e potencialmente governadas por regras” (Kahneman, 2003, p.698, tradução nossa). Eis alguns exemplos das atividades do Sistema 2, em ordem de complexidade:

- Manter-se no lugar para o tiro de largada numa corrida.
- Concentrar a atenção nos palhaços do circo.
- Concentrar-se na voz de determinada pessoa em uma sala cheia e barulhenta.
- Procurar uma mulher de cabelos brancos.
- Sondar a memória para identificar um som surpreendente.
- Manter uma velocidade de caminhada mais rápida do que o natural para você.
- Monitorar a conveniência de seu comportamento numa situação social.
- Contar as ocorrências da letra “a” numa página de texto.
- Dizer a alguém seu número de telefone.
- Estacionar numa vaga apertada (para a maioria das pessoas, exceto manobristas de garagem).
- Comparar duas máquinas de lavar roupa em relação ao valor global.
- Preencher um formulário de imposto.
- Verificar a validade de um argumento lógico complexo. (KAHNEMAN, 2012, p.31)

Ademais, há uma interação constante entre os sistemas. Por exemplo, quando nos deparamos com algo que requer nossa atenção, julgamentos rápidos são gerados pelo Sistema 1. Posteriormente, se necessário, esses julgamentos são avaliados e corrigidos pelo Sistema 2. Entretanto, conforme observamos no uso inadequado da heurística da representatividade, a interação nem sempre ocorre de maneira eficaz, resultando em situações em que um julgamento equivocado pode prevalecer.

Como o Sistema 1 opera automaticamente e não pode ser desligado a seu bel-prazer, erros do pensamento intuitivo muitas vezes são difíceis de prevenir. Os vieses nem sempre podem ser evitados, pois o Sistema 2 talvez não ofereça pista alguma sobre o erro. Mesmo quando dicas para prováveis erros estão disponíveis, estes só podem ser prevenidos por meio do monitoramento acentuado e da atividade

diligente do Sistema 2. Como um modo de viver sua vida, porém, vigilância contínua não necessariamente é um bem, e certamente é algo impraticável. Questionar constantemente nosso próprio pensamento seria impossivelmente tedioso, e o Sistema 2 é vagaroso e ineficiente demais para servir como um substituto para o Sistema 1 na tomada de decisões rotineiras. (KAHNEMAN, 2012, p.38–39)

Distinta da visão propiciada pela teoria do processo dual, existe ainda todo um corpo de estudos que concentra suas pesquisas na interação entre os agentes e seu ambiente circundante. Uma vez que é praticamente impossível prospectarmos todas as informações ambientais disponíveis, pesquisadores como Fiedler e Juslin (2006) sustentam que, em relação aos dados que alimentam nossas inferências rotineiras, nos comportamos frequentemente como “estatísticos ingênuos”. Ou seja, independentemente de sua origem ou consistência, aceitamos as amostras de informação que nos são fornecidas sem maiores questionamentos. Portanto, a falha de julgamento não seria ocasionada pela maneira como as operações cognitivas do agente se interrelacionam, conforme advoga a teoria do processo dual, mas sim pelo processamento de uma informação tendenciosa.

A noção de um estatístico intuitivo ingênuo pode de fato ser uma explicação de que a mente se aproxima de um estatístico propriamente dito. Consistente com o programa de heurísticas e vieses, algumas das falhas para otimizar as estimativas de uma amostra podem derivar de um processamento cognitivo restrito e heurístico. [Contudo,] muitos dos fenômenos que tradicionalmente foram explicados em termos, digamos, da heurística de disponibilidade [...] podem com igual força ser interpretados como resultados de um processamento ótimo por contagens de frequência que se baseiam em amostras tendenciosas. Por exemplo, equívocos sobre o risco de várias causas de morte podem, em parte, derivar de vieses na cobertura jornalística de acidentes e doenças, ou seja, a amostra ambiental é o que está influenciando o julgador, em vez do processamento tendencioso dessa entrada. (FIEDLER; JUSLIN, 2006, p.6, tradução nossa)

Por outro lado, Le Mens e Denrell (2011) argumentam que independentemente da ingenuidade ou da sagacidade do agente, o próprio modo como conduzimos nossas avaliações racionais poderia ser a causa de muitos resultados inadequados.

Explicações por amostragem de julgamentos aparentemente tendenciosos não precisam depender de uma suposição de “ingenuidade”. Não é necessário supor que os tomadores de decisão não estejam cientes dos vieses nas amostras de informações que usam para formar julgamentos. Em vez disso, erros sistemáticos de julgamento podem surgir apenas devido a assimetrias nas informações disponíveis para o tomador de decisão quando seu objetivo é maximizar seu retorno esperado. De maneira mais geral, mostrar que a suposição de ingenuidade não é necessária para explicações de amostragem expande o escopo e o alcance da aplicabilidade dessa classe de explicações. (LE MENS; DENRELL, 2011, p.388, tradução nossa)

Essa linha de estudos defende que o objetivo de maximizar os resultados, um aspecto inerente aos nossos processos de decisão racional, acarreta certas imposições que podem gerar paradoxalmente resultados enviesados. Uma vez que a maioria dos processos que executamos no nosso dia a dia é de natureza indutiva⁷, não é contrário à razão preferirmos informações incidentais, facilmente obtidas no nosso meio social, a informações ambientais menos acessíveis e escassas. Assim, a disponibilidade assimétrica das informações, efeito da inevitável socialização, bem como a própria dinâmica de nossos processos racionais, poderia explicar muitos dos nossos erros de julgamento.

[...] a tendência de favorecer as alternativas para as quais a informação é mais acessível, pode ser consistente com o comportamento racional. [O modelo que simula nossos processos racionais de decisão] oferece uma explicação racional para o comportamento que antes era atribuído a falhas no processamento e no cálculo das informações, como a tendência de se ter impressões mais positivas daquelas pessoas frequentemente mais à vista. (LE MENS; DENRELL, 2011, p.380, tradução nossa)

Em suma, de acordo com as abordagens da psicologia cognitiva, possíveis erros de julgamento podem ser explicados por limitações do próprio agente, tanto no que diz respeito aos vieses cognitivos e às heurísticas a que está sujeito, quanto à aceitação acrítica de amostras informacionais ambientais; ou ainda, por limitações suscitadas pela própria dinâmica dos processos decisórios em sociedade. Vale observar como esses aspectos, fixados segundo as visões da psicologia cognitiva, ao

⁷ Em um processo indutivo, quando estamos diante de duas alternativas possíveis, a razão determina que a escolha recaia sobre aquela que é sustentada por uma maior quantidade de informações. No entanto, isso não invalida o resultado rejeitado. Este poderia ser reconsiderado, caso novas informações viessem à luz.

determinar o comportamento do agente, acabam por afastá-lo do caráter racional prescrito pelo modelo do *homo economicus*.

Dessa forma, poder-se-ia compreender o porquê da ocorrência do fenômeno da polarização: quando agentes deliberam em grupo, os erros individuais de julgamento, explicados segundo as formulações estudadas acima, se somam, conduzindo todo o grupo a extremos de crenças e opiniões.

Uma vez que temos noção dos aspectos responsáveis por conduzir um agente e, por conseguinte, todo um grupo ao erro, não poderíamos de alguma forma minorar ou mesmo eliminar alguns dos seus efeitos? Em tese, a educação, o conhecimento adquirido, o treino intelectual e o consequente desenvolvimento de novas ferramentas de análise permitiriam tanto uma melhor avaliação dos nossos erros de julgamento quanto o desenvolvimento de procedimentos de correção. Ora, não seria essa a ideia básica que permeia as especulações filosóficas desde Sócrates e Platão? O exercício da razão e a busca constante pelo conhecimento não seriam os remédios mais eficazes contra as possíveis falhas ético-morais ou cognitivas que pudéssemos apresentar?

Mais precisamente, no caso da teoria do processo dual, uma constante análise dos erros e dos acertos, o desenvolvimento de novas ferramentas analíticas e tecnológicas,⁸ bem como o conhecimento adquirido, seriam capazes, pelo menos em tese, de mitigar e até mesmo evitar a ação de alguns vieses cognitivos.⁹ Igualmente, essas ações poderiam aprimorar nossos procedimentos de amostragem, tornando-nos amostradores mais precavidos tanto em relação às falhas e às assimetrias das informações disponíveis quanto aos eventuais desdobramentos e desvios a que estariam sujeitos nossos processos decisórios em sociedade. Por exemplo, a

⁸ Os sistemas atuais de inteligência artificial seriam um bom exemplo de ferramentas desse tipo.

⁹ Em um exemplo ilustrativo desse fato, Kahneman (2012) cita a famosa ilusão de Müller-Lyer. Essa ilusão consiste em um conjunto de linhas que terminam em pontas de seta. Embora os conjuntos de linhas apresentem o mesmo tamanho, a orientação das pontas de cada seta determina o modo como percebemos o comprimento das linhas. As linhas que apresentam os ângulos das pontas voltados para dentro parecem ter um comprimento menor que as outras, cujas pontas apresentam ângulo voltado para fora. Nessa ilusão, muito conhecida, ainda que o sistema 1 nos indique o oposto, as operações do sistema 2 (devido ao nosso conhecimento prévio sobre o fato de que se trata de uma ilusão) sugerem que todas as linhas apresentam o mesmo tamanho.

ortodoxia econômica adota praticamente o modelo racional do *homo economicus* como um axioma. Contudo, seus representantes, no sentido de contornarem as falhas e imprecisões que surgiram principalmente nas aplicações matemáticas de suas formulações, desenvolveram procedimentos de correção, como os denominados “termos de erro”:

Economistas se sentiam à vontade com a ideia de que seus modelos eram imprecisos e que as previsões decorrentes desses modelos continham falhas. Nos modelos estatísticos usados pelos economistas, lida-se com isso simplesmente introduzindo na equação um “termo de erro”. Suponha que, para prever a altura que uma criança atingirá na idade adulta, você use a altura de ambos os pais como preditores. Esse modelo fará um serviço decente, já que pais altos tendem a ter filhos altos, mas não será perfeitamente acurado, e é isto que o termo de erro procura capturar. E enquanto os erros forem aleatórios — isto é, as previsões do modelo forem altas ou baixas demais com igual frequência —, está tudo bem. Os erros se anulam mutuamente. Era este o raciocínio dos economistas para justificar por que os erros produzidos pela racionalidade limitada podiam ser ignorados com segurança. E voltamos ao modelo totalmente racional! (THALER, 2019, p.37–38)

Tendo isso em vista, podemos pensar que pessoas mais preparadas do ponto de vista intelectual, bem como indivíduos dotados de melhores instrumentos analíticos, seriam menos sensíveis àqueles aspectos limitadores da ação racional individual, conforme asseveram as diversas visões da psicologia cognitiva. O que nos levaria a inferir que um grupo composto por indivíduos mais preparados do ponto de vista intelectual dificilmente se polarizaria. Entretanto, em determinadas situações, o desenvolvimento de novos instrumentos de análise, a educação e o conhecimento adquirido nem sempre surtem os efeitos desejados. Por exemplo, hoje há praticamente consenso na comunidade científica que o aquecimento global representa uma grave ameaça à continuidade da nossa existência. Contudo, é notória também a percepção disseminada entre o público leigo que tal temor é exagerado e até sem sentido. Essa percepção equivocada poderia ser explicada particularmente por uma distribuição desigual do conhecimento científico-matemático entre a população. Dessa forma, indivíduos com melhor treinamento intelectual nessas áreas estariam mais habilitados a avaliar os graves riscos envolvidos no aquecimento global do que aqueles com menos treino. Kahan *et al.* (2012) estudaram e testaram esse

argumento, analisando a percepção de risco das mudanças climáticas globais em uma amostra de 1540 indivíduos, e os resultados foram surpreendentes. A percepção de risco apresentou-se negativamente correlacionada com o preparo científico-matemático dos agentes. Ou seja, indivíduos mais preparados apresentaram uma menor percepção de risco às mudanças climáticas globais. Contudo, ao se decompor a amostra do estudo em duas classes, a saber, uma de indivíduos de perfil conservador¹⁰ e outra de perfil progressista¹¹, notou-se uma diferença significativa em relação às percepções de risco. Assim, quando a percepção de risco foi avaliada por indivíduos de perfil conservador, além de se apresentar negativamente correlacionada ao respectivo preparo, exibiu uma taxa de variação bastante acentuada entre as grandezas envolvidas. Ou seja, nesse caso, um maior preparo resultou em uma acentuada piora na avaliação de risco associado às mudanças climáticas. Vale notar que, nesse caso específico, os conservadores com menos preparo já apresentavam avaliações negativas em relação ao risco. Por outro lado, entre indivíduos de perfil progressista, a percepção de risco apresentou uma correlação positiva com o preparo, embora a taxa de variação tenha sido pequena. Isto é, nesse caso, um maior preparo resultou em um aumento modesto na avaliação de risco associado às mudanças climáticas. Vale notar que, nesse caso específico, os progressistas com menos preparo já apresentavam avaliações positivas em relação ao risco.

Segundo os autores do artigo, as especificidades dos grupos de pertença dos agentes estariam na raiz de tal fenômeno:

Dado que o indivíduo comum depende de seus pares — material e emocionalmente — e o pouco impacto das suas crenças no seu ambiente físico, provavelmente seria melhor para ele formar

¹⁰ Segundo os autores, “indivíduos que possuem essa visão de mundo hierárquica e individualista — [visão] que associa o poder da autoridade a claros rankings sociais, acabando por desconsiderar qualquer interferência coletiva nas decisões dos indivíduos que detêm tal ascendência — tendem a ser céticos em relação aos riscos ambientais. Essas pessoas percebem intuitivamente que a aceitação generalizada desses riscos legitimaria restrições à prática do comércio e da indústria, formas de comportamento que os individualistas hierárquicos [conservadores] valorizam”. (KAHAN *et al.*, 2012, p.732, tradução nossa)

¹¹ Segundo os autores, “indivíduos que possuem uma visão de mundo igualitária e comunitária — [visão] que favorece formas de organização social menos regimentadas e maior atenção coletiva às necessidades individuais — tendem a ser moralmente desconfiadas do comércio e da indústria, aos quais atribuem a desigualdade social. Por isso, são propensos a acreditar que tais comportamentos são perigosos e dignos de restrição. Nessa perspectiva, esperaríamos que os comunitários igualitários [progressistas] estivessem mais preocupados com os riscos das mudanças climáticas do que os individualistas hierárquicos [conservadores]”. (KAHAN *et al.*, 2012, p.732–733, tradução nossa)

percepções de risco que minimizassem qualquer perigo de afastamento da sua comunidade. (KAHAN *et al.*, 2012, p.734, tradução nossa)

Ademais, por que os maiores valores das avaliações de risco, a despeito de serem positivas ou negativas, foram apresentadas por progressistas ou por conservadores mais bem preparados tecnicamente?

Para os cidadãos comuns, a recompensa por adquirir um maior conhecimento científico e uma capacidade mais confiável de raciocínio técnico é uma maior facilidade em descobrir, usar — ou explicar — as evidências relacionadas às posições dos seus grupos. (KAHAN *et al.*, 2012, p.734, tradução nossa)

Dessa forma, podemos pensar que as especificidades de seus grupos de pertencimento, em determinadas ocasiões, dizem muito mais sobre como os indivíduos se posicionam em sociedade que outros aspectos mais visíveis, como o nível de instrução ou o conhecimento técnico, por exemplo.¹² Tal situação não deixa de ser preocupante, uma vez que julgamentos gerados nessas circunstâncias, em que há o claro envolvimento de grupos de pertença, podem ser muito prejudiciais ao bem-estar comum.

Mesmo que a cognição cultural sirva aos interesses pessoais dos indivíduos, essa forma de raciocínio pode ter um impacto altamente negativo na tomada de decisões coletivas. O que orienta a percepção individual de risco, nesse sentido, não é a verdade dessas crenças, mas sim a sua congruência com os compromissos culturais dos indivíduos. [...] Embora seja efetivamente sem custo para qualquer indivíduo formar uma percepção de risco de mudança climática equivocada, mas culturalmente agradável, é muito prejudicial ao bem-estar da coletividade a agregação de crenças geradas dessa maneira. (KAHAN *et al.*, 2012, p.734, tradução nossa)

O estudo de caso que acabamos de apresentar confirma um fato — concepções fundamentais arraigadas em nosso senso comum, algumas vezes, não se mostram corretas. Este é o caso da ideia de que indivíduos com maior nível

¹² Abordaremos essa mesma questão adiante; contudo, sob uma perspectiva epistemológica.

educacional e técnico, independentemente das circunstâncias, realizariam avaliações mais críticas e apuradas na maior parte do tempo. Como vimos, em determinadas situações, isso é no mínimo discutível. Ademais, quando indivíduos com bom treino intelectual apresentam posições questionáveis a respeito de conhecimentos praticamente consolidados a que teriam fácil acesso, isso sugere que seus respectivos grupos de pertença poderiam estar polarizados em relação às crenças conexas a esses assuntos. Assim, a notória importância que esse estudo de caso atribui à influência dos contextos coletivos no comportamento individual, pode ser uma pista valiosa para um melhor entendimento da ocorrência do fenômeno da polarização. Compreender agora, de modo mais preciso, como se processa essa influência coletiva, exige o estudo de novos instrumentos analíticos, como veremos a seguir.

2.1 A lógica da polarização grupal

Corroborados por dados experimentais, estudos atuais no campo da psicologia social mostram-nos a ocorrência de muitos fenômenos que se originam notadamente quando indivíduos interagem no interior de um grupo. Esses estudos procuram explicar tais fenômenos, centrando a atenção no papel de determinados processos deliberativos na construção de convicções e de consensos grupais. Entre os fenômenos examinados nesse campo de estudos,¹³ o da polarização grupal se destaca:

[Durante] grande parte do tempo, grupos de pessoas acabam pensando e fazendo coisas que os membros do grupo nunca pensariam ou fariam sozinhos. Isso vale para grupos de adolescentes dispostos a correr riscos que individualmente evitariam. Certamente vale para aqueles inclinados à violência, inclusive terroristas e os que cometem genocídio. Vale para investidores e executivos. Vale para oficiais do governo, associações de bairro, reformadores sociais, contestadores políticos, policiais, organizações estudantis, sindicatos trabalhistas e júrís. Alguns dos melhores e dos piores desdobramentos da vida social são produtos da dinâmica de grupo, na qual os membros das organizações, grandes e pequenas, se movem em novas direções. [...] O tema unificador é simples: quando *as pessoas se*

¹³ Além da polarização, outros fenômenos grupais estudados pelo campo da psicologia social são: a facilitação social, a vadiagem social, a desindividuação, o pensamento de grupo e a influência da minoria.

encontram em grupos de tipos que pensam de modo semelhante, ficam particularmente propensas a se mover para os extremos. E quando tais grupos incluem autoridades que dizem aos membros do grupo o que fazer ou que conferem a eles certos papéis sociais, muitas coisas ruins podem acontecer. (SUNSTEIN, 2010, p.4, grifo do autor)

Por exemplo, de acordo com essa formulação, em uma deliberação financeira entre indivíduos com propensão ao risco, a decisão final quase sempre se apresenta mais arriscada que a inclinação inicial dos agentes. O oposto também ocorre, ou seja, quando em deliberação, indivíduos com temor ante à possibilidade de perdas geralmente apresentam decisões grupais mais conservadoras. Da mesma forma, se um grupo de altruístas em deliberação decidir ajudar uma instituição de caridade, observaremos doações mais generosas do que aquelas a que esses indivíduos estariam inicialmente dispostos a fazer. Em suma, na polarização, as opiniões movimentam-se sempre no sentido da inclinação comportamental inicial da maioria dos agentes¹⁴. Se tal inclinação se mostrar preconceituosa, pouco informada ou até mesmo irracional, a deliberação grupal não consegue reverter tal tendência. Aliás, ao contrário, ela é o principal indutor da polarização nesses casos.

2.2 As causas da polarização grupal

Entre as muitas teorias do campo da psicologia social que pretendem explicar o porquê da ocorrência da polarização, segundo Myers (2014), duas se destacam: **a influência informacional** (ou a interpretação da persuasão argumentativa) e **a influência normativa** (ou a interpretação da comparação social). Em uma dada deliberação, segundo **a teoria da influência informacional** (a teoria mais bem sustentada pelos estudiosos do tema), a exposição a novas informações e a inéditos argumentos e explicações, desde que em consonância com o ponto de vista dominante, afetariam toda a estrutura discursiva grupal, polarizando os discursos compartilhados no interior do grupo. Ademais, há certos comportamentos que acabam por intensificar esse processo de polarização:

A participação ativa na discussão produz mais mudança de atitude do que a escuta passiva. Participantes e observadores ouvem as mesmas ideias, mas quando os participantes as expressam em suas

¹⁴ Grupos formados por indivíduos com diferentes tendências de pensamento dificilmente polarizam.

próprias palavras, o compromisso verbal amplia o impacto. Quanto mais os membros do grupo repetem as ideias um do outro, mais eles as ensaiam e validam. (MYERS, 2014, p. 231)

Vale ressaltar um aspecto importante nessa teoria:

Os que carecem de confiança e que não têm certeza do que devem pensar tendem a conter suas opiniões. Suponha que lhe perguntem o que você acha sobre algumas questões para as quais não tem informação. Provavelmente, você evitará os extremos. É por isso que as pessoas cautelosas, não sabendo o que fazer, tendem a escolher algum ponto médio entre os extremos. Mas se as outras pessoas parecem compartilhar de suas concepções, as pessoas se tornam mais confiantes de que estão certas. Como resultado, provavelmente mudarão para uma direção mais extrema. [...] Em uma grande variedade de contextos experimentais, as opiniões das pessoas têm se mostrado mais extremas porque suas concepções iniciais foram corroboradas e porque ficariam mais confiantes depois de conhecerem as concepções compartilhadas por outras. (SUNSTEIN, 2010, p.25)

Por seu turno, **a teoria da influência normativa** propõe a ideia de que estamos constantemente nos comparando aos outros e, por essa razão, procuramos apresentar opiniões e ideias acima da média quando estamos em grupo. Ao ser compartilhada simultaneamente por todos os integrantes de um determinado grupo, tal comportamento conduziria inevitavelmente à polarização. Exemplificando, se em um grupo de liberais o melhor é ser mais liberal, ou em um grupo de conservadores, mais conservador, a consequência natural desse processo de comparação social seria a polarização das posições. A divertida situação, narrada por Sunstein, ilustra bem essa teoria:

Há alguns anos, eu estava discutindo polarização grupal com um filósofo que trabalha com direitos dos animais e bem-estar animal. Ele é muito dedicado a reduzir o sofrimento dos animais e me contou a seguinte história: “Em uma sexta-feira de uma conferência de três dias, estávamos perfeitamente sensatos, a meu modo de ver. Mas quando chegou o domingo, paramos de pensar direito! Ficamos muito mais extremistas. No domingo, as pessoas começaram a dizer que nenhum experimento em animais jamais produziu algum conhecimento útil aos seres humanos. No domingo, as pessoas começaram a dizer que nunca é aceitável comer carne, mesmo que os animais tenham tido uma vida longa e feliz e tenham morrido de causas naturais. De algum modo, perdemos a cabeça. O filósofo me disse que essa mudança de perspectiva — uma forma de polarização - não era adequadamente explicada pela troca de novas informações ou pelo aumento da confiança. (SUNSTEIN, 2010, p. 27)

2.3 A polarização grupal e as bolhas epistêmicas

Alguns grupos polarizados nitidamente possuem uma duração maior que outros. Há grupos efêmeros, como uma assembleia sindical ou uma reunião de moradores que podem apresentar momentaneamente decisões mais extremadas. E há ainda aqueles grupos que persistem por mais tempo. Por exemplo, um agrupamento de pessoas ideologicamente constituído como um grupo terrorista ou ainda um grupo fechado de pessoas no *WhatsApp* que defende teses terraplanistas. Vale notar que essa persistência pode durar meses e até anos!

Ao procurar figurar o modo de apresentação e a organização das estruturas socio-epistêmicas de um grupo polarizado mais persistente, a literatura sobre o tema utiliza os seguintes termos: “câmara de eco”, “bolha epistêmica” e “bolha-filtro”. Especialmente, este último termo tem sido utilizado com frequência nos estudos do fenômeno da polarização nas mídias sociais, para caracterizar aqueles grupos virtuais que supostamente se polarizam devido à atuação dos denominados “algoritmos curadores”, como veremos em maior detalhe adiante.

De modo geral, ainda de acordo com a literatura, pode-se utilizar indistintamente os termos “câmara de eco” e “bolha epistêmica” na caracterização de um grupo polarizado persistente. Vale notar ainda que esses termos também podem ser utilizados na caracterização dos grupos virtuais.

De qualquer modo, na direção de uma definição mais precisa, de acordo com Lackey (2021a) e Robson (2014), podemos dizer que um determinado grupo de pessoas se caracteriza como uma **câmara de eco** ou **bolha epistêmica** quando:

- 1) a visão aceita pelo grupo (particularmente dos líderes de opinião) é amiúde reiterada e reforçada, enquanto opiniões discordantes, caso ocorram, são sufocadas ou desconsideradas;
- 2) de algum modo, são espaços – sejam geográficos, culturais, ou em qualquer outro aspecto em que possam se apresentar –, fechados;
- 3) os julgamentos dos formadores de opinião não são meramente transmitidos, são também amplificados.

Por seu turno, em um sentido mais formal, temos a definição de Woods (2005, p.740, tradução nossa, negrito nosso), “[um] agente cognitivo **X** ocupa **uma bolha epistêmica** precisamente quando é incapaz de distinguir entre o seu pensamento de que conhece ‘**p**’ e o seu conhecimento de ‘**p**’”. Esclarecendo. Se partimos da definição clássica de conhecimento como crença verdadeira e justificada, a definição de Woods destaca a condição na qual um indivíduo no interior de uma bolha não consegue distinguir entre o que acredita conhecer e o que efetivamente conhece. Assim, em uma bolha epistêmica, o agente encontra-se praticamente confinado em seu próprio sistema de crenças e justificações, privado do benefício de perspectivas externas que poderiam de alguma forma questionar ou validar suas convicções. Portanto, no interior de uma bolha, embora se possa ter a impressão de conhecimento, há o risco de não se alcançar o conhecimento genuíno, pois as justificativas que sustentam as crenças não são disputadas nem corroboradas adequadamente.

Por fim, vale observar ainda a definição de Anderson (2021) que parece fazer uma síntese entre as formulações de Lackey e de Woods, destacando, contudo, principalmente a dificuldade de correção de um raciocínio falacioso no interior de uma bolha:

Uma bolha epistêmica pode ser definida como uma rede social relativamente autosegregada de pessoas que pensam de modo semelhante, que carece de disposições internas para desacreditar afirmações factuais falsas ou não comprovadas em domínios específicos. Devido a fatores internos à rede, os membros são propensos a convergir e resistir à correção de afirmações falsas, enganosas ou não comprovadas que circulam no seu interior. (Anderson, 2021, p.11, tradução nossa, negrito nosso)

Na tentativa de estabelecer uma terminologia mais precisa, Nguyen (2018) advoga a existência de uma diferença sutil principalmente entre os termos “bolha epistêmica” e “câmara de eco”. Para o autor, estaríamos diante de uma bolha epistêmica quando opiniões discordantes à tendência grupal fossem desconsideradas no seu interior. No entanto, caso estivéssemos diante de uma câmara de eco, essas opiniões seriam ativamente sufocadas. Nesse sentido, uma câmara de eco é um fenômeno bem mais complexo e danoso à sociedade que uma bolha epistêmica. Entretanto, em muitos casos, não é uma tarefa simples detectar se um grupo polarizado somente desconsidera ou sufoca as opiniões divergentes. Desse modo, no sentido de se evitar que disputas terminológicas dificultem desnecessariamente nosso

estudo, utilizaremos doravante somente o termo “bolha epistêmica” quando quisermos caracterizar de modo mais preciso os grupos polarizados persistentes. Além disso, por expressarem faces distintas de um mesmo fenômeno, consideraremos os elementos que compõem as definições de Lackey (2021a), Woods (2005) e Anderson (2021) como complementares.

3 EXEMPLOS DE BOLHAS EPISTÊMICAS

Para compreendermos o funcionamento de uma bolha epistêmica na prática, apresentaremos brevemente três exemplos ilustrativos. O primeiro, um dos momentos mais infames da história. O segundo, embora prosaico, dadas suas características peculiares, não deixa de ser um exemplo valioso. Por fim, o terceiro, acreditamos, configura-se como uma das maiores ameaças ao nosso atual arcabouço civilizatório.

3.1 A doutrina da bruxaria no período renascentista

Por volta da segunda metade do século XVI, a elite letrada europeia se polarizou em torno de alguns temas, alcançando praticamente o estado de uma bolha epistêmica de efeitos trágicos. Um importante autor renascentista concentrou sua atenção no discurso explicativo que se originara a partir de tal estado. Na obra “Os Ensaios”, mais precisamente, no ensaio intitulado “Dos Coxos”¹⁵, Michel de Montaigne, entre outras especulações, apresenta-nos um quadro bastante vívido das discussões corriqueiras entabuladas pela elite francesa nos jantares, encontros e reuniões no final do século XVI. Temas como a bruxaria e seus correlatos (milagres extraordinários, possessões demoníacas, estranhos curandeiros) eram discutidas com aparente rigor intelectual. Montaigne identificou precisamente no mau uso da razão, particularmente nas construções explicativas falaciosas, um sinal preocupante desse estado de coisas:

Estava a meditar, há pouco, como faço amiúde, sobre a vagueza e a disponibilidade desse instrumento mal regulado que é a razão humana. Vejo comumente que os homens, em lugar de atentar para a realidade dos fatos, se divertem com lhes buscar as causas. Passam por cima dos antecedentes e atêm-se a dominar minuciosamente as conseqüências. Deixam as coisas e correm às causas. (MONTAIGNE, 2016, p.946)

Ainda, segundo o filósofo, pressupostos básicos são ignorados em muitas situações, como a própria ocorrência dos eventos que se quer explicar, por exemplo:

¹⁵ O termo “coxo” é utilizado aqui em seu sentido histórico, referindo-se a pessoas com alguma limitação física que afeta sua locomoção. Estamos cientes e reconhecemos que este termo pode ser considerado capacitista sob as normas recentes de utilização da linguagem. Entretanto, optamos por mantê-lo com o objetivo de preservar a intenção original do filósofo.

Ao ouvir falar de alguma coisa, começa-se por indagar: "como é?". Dever-se-ia dizer: "antes de mais nada, é?". Nosso raciocínio é capaz de reconstruir um mundo como o nosso e descobrir-lhe os princípios e a organização; não precisa para tanto nem de base nem de materiais; basta-lhe deixar-se levar, "hábil que é em dar um corpo à fumaça". Constrói tão bem sobre o vazio como sobre o cheio, com nada como com alguma coisa. (MONTAIGNE, 2016, p.947)

Como sabemos, o ponto de partida de qualquer explicação deveria ser sempre a verificação do fato que se pretende explicar. Somente após a constatação da sua ocorrência é que toda a maquinaria explanatória deveria entrar em cena, concentrando-se nos possíveis encadeamentos causais necessários para a sua explicação. Como tal estado de coisas não era percebido pelos ilustres debatedores à mesa do jantar?

Acho que de quase tudo deveríamos dizer: "isto não é". Daria amiúde uma tal resposta se ousasse; mas logo proclamam que falar dessa maneira denota ignorância e fraqueza de espírito e tenho, a maior parte do tempo, que representar em companhia dos outros e conversar sobre assuntos frívolos a que não dou fé. Sem contar que é em verdade algo grosseiro e peculiar ao espírito de contradição negar categoricamente um fato que nos afirmam. Tanto mais quanto poucas pessoas deixam de insistir em que o viram, indicando testemunhas autorizadas; principalmente quando o fato é pouco digno de crédito. (MONTAIGNE, 2016, p.947)

[...] Bem vejo que se irritam e me proibem a dúvida sob pena das piores injúrias; é um novo método de persuasão. Mas graças a Deus não será a socos que me hão de impor uma orientação. Compreendo que aqueles cuja opinião é tachada de falsidade se revoltam contra a apreciação. (MONTAIGNE, 2016, p.951)

O filósofo alerta-nos ainda para a ocorrência de um segundo problema: a pressão social exercida sobre todos aqueles que de alguma forma contrariavam a tendência de pensamento do grupo. Nesse aspecto, Montaigne nos dá a entender que cedeu à pressão em muitas oportunidades. Lembremos que o autor era também um nobre e por mais ciente que estivesse das explicações falaciosas que circulavam no interior de seu grupo, procurava não se indispor publicamente.

Havia ainda um terceiro conjunto de problemas: a existência de uma classe de teorias e crenças compartilhadas no interior de seu grupo que estreitavam a margem de manobra discursiva em relação a determinados temas. Entre 1584 e 1585, enquanto Montaigne finalizava o ensaio citado, um determinado esquema conceitual

alcançava seu ápice. Esse esquema proporcionou às elites culturais europeias, que há cem anos exaltavam o renascimento do ideal clássico, a possibilidade de disseminar e sustentar uma das mais sórdidas mistificações da história: a existência das bruxas.¹⁶ E quais seriam as causas dessa aparente mistura, a princípio tão incoerente, de crenças humanísticas e místicas à época da Renascença? Para respondermos a essa questão, precisamos detalhar o imaginário renascentista. Nesse sentido, esclarece Koyré (1982):

O tipo que encarna o ambiente e o espírito da Renascença é, evidentemente, o grande artista. Mas é também, e talvez principalmente, o homem de letras. Foram os homens de letras os seus promotores, os seus anunciadores e “trombeteadores”. Foram também os eruditos. E aqui, permito-me lembrar-lhes o que disse Bréhier: o espírito de erudição não coincide inteiramente — e até não coincide de maneira alguma — com o espírito da ciência.

Por outro lado, sabemos também — e isto é muito importante — que **a época da Renascença foi uma das épocas menos dotadas de espírito crítico que o mundo conheceu. Trata-se da época da mais grosseira e mais profunda superstição, da época em que a crença na magia e na feitiçaria se expandiu de modo prodigioso, infinitamente mais do que na Idade Média.**

[...] E se examinarmos a produção literária dessa época, tornar-se-á evidente que não são os belos volumes das traduções dos clássicos produzidos nas tipografias venezianas que constituem os grandes sucessos de livraria; são as demonologias e os livros de magia. (KOYRÉ, 1982, p. 46–47, negrito nosso)

Vale notar que, segundo o texto de Koyré, a existência de um mercado editorial já era uma realidade naquela época. Em torno do ano de 1450, a invenção da prensa de tipos móveis por Johannes Gutenberg possibilitou o surgimento da primeira mídia de massa da história — o livro impresso. Em uma escala até então inimaginável, a impressão em série de livros propiciou a rápida circulação de ideias pelo continente europeu. Esse fato inédito, aliado a uma estranha mistura entre o gosto da elite renascentista pela literatura profana e mística e a prática de uma fervorosa religiosidade, fez crescer o medo de que estivesse em curso uma autêntica conspiração demoníaca contra toda a cristandade, justificando as crenças mais delirantes. Entre elas, destacava-se a crença na existência da bruxaria que, segundo Clark (2006), alcançou praticamente um estatuto epistemológico.

¹⁶ Estima-se que no período entre o início da Renascença e o início do Iluminismo, de 40.000 a 60.000 pessoas tenham sido condenadas e executadas pelo crime de bruxaria. Foi o período de maior perseguição à bruxaria na história.

Dado que a existência da bruxaria, a partir de determinado momento, ultrapassara a mera credulidade, duas questões se apresentam: i) já conhecemos o seu veículo de propagação, o livro impresso, mas qual seria a gênese desse conjunto de crenças? ii) Por que, conforme afirma Koyré, o raciocínio crítico apresentava-se tão fragilizado naquela época?

Quanto à primeira questão, há consenso em se afirmar que, a partir de costumes, crenças e teorias cujas origens remontavam à Baixa Idade Média, um determinado conjunto de ideias condensara-se numa autêntica **doutrina da bruxaria** entre o final do século XV e meados do século XVI. Como afirma Ostorero (2020), as principais características desse conjunto de crenças eram:

- uma forte dimensão coletiva que se expressava numa espécie de seita ou sociedade secreta;
- a clara adoração ao demônio ou apostasia, em que se estabelecia um vínculo de fidelidade absoluta que unia a bruxa ao diabo;
- os atos dessa sociedade secreta instauravam-se em oposição à natureza: seus integrantes praticavam o canibalismo, cometiam infanticídio e realizavam atos sexuais desviantes;
- As bruxas podiam voar.

Chamava-se “sabá” a cerimônia secreta em que todas essas características confluíam de algum modo. Esse ritual era descrito nos seus pormenores fantásticos, independentemente da região ou do idioma, por praticamente todos os sentenciados¹⁷ pelo crime de bruxaria.

Uma vez que a doutrina da bruxaria foi uma edificação erguida no seio da elite renascentista, não é surpreendente que filósofos, teólogos e juristas tenham sido seus principais arquitetos. Estes, à medida que a doutrina se consolidava, desenvolveram um rigoroso arcabouço teológico e jurídico. Quanto a este aspecto jurídico, vale notar que os julgamentos eram realizados tanto por autoridades eclesiásticas quanto seculares. Normalmente, os réus eram acusados de *maleficia*, encantamento mágico-demoníaco que possuía algum objetivo nefasto, como a perda de alguma safra, a

¹⁷ A maioria dos réus era composta por mulheres. Havia um caráter fortemente misógino nessa perseguição.

doença de alguma pessoa, a morte de algum animal ou o mau tempo. O mesmo ritual era seguido em todos os lugares. Na maioria das vezes, um aldeão acusava uma viúva, um vizinho ou uma mulher idosa. Os depoimentos eram tomados após a tortura dos réus ou mesmo na sua mera ameaça. As vítimas diziam o que os inquisidores queriam ouvir e eram frequentemente instadas a denunciar seus vizinhos, em uma progressão inimaginável de terror e morte.

Neste ponto, seria importante refletirmos sobre a seguinte questão: sugerir, como fizemos, que principalmente o medo de uma conspiração demoníaca fosse capaz de proporcionar a formação da doutrina da bruxaria, bem como de toda uma estrutura teológica-jurídica de repressão, não seria uma visão simplista do fenômeno? Não existiriam fatores econômicos ou sociais que também poderiam explicar essa onda de perseguições? Segundo Clark (2006), Delumeau (2009) e Levack (1988), entre os séculos XV e XVII, a Europa presenciou inúmeras mudanças sociais, políticas, econômicas e religiosas. Contudo, um fato diferenciava esse período de outros igualmente marcantes: a intensidade dos acontecimentos combinada com a enorme rapidez com que tudo ocorria.¹⁸ Isso definitivamente contribuía para a instauração da incerteza e do medo, percepções que colocavam em risco a ordem religiosa e por tabela toda a estrutura social. Tal situação acabava por sitiar a elite dirigente, detentora dos principais instrumentos de controle social, conduzindo-a inevitavelmente à violenta repressão contra a bruxaria, um dos principais símbolos desse estado conflituoso.

As razões pelas quais as bruxas foram responsabilizadas pela desordem, ou consideradas sintomáticas dela, tinham, contudo, a ver com crenças; ou, se preferirmos, com suposições, atitudes, medos, preconceitos, ou o que quer que seja. E só poderemos chegar a essas coisas, insisto, se às percepções dos envolvidos for conferida a necessária autonomia referencial. Os ataques à bruxaria não foram produtos acidentais do trauma social, produtos que, de um ponto de vista causal, poderiam ter tomado uma forma bem diferente. Pelo contrário, a própria noção de bruxaria era inseparável de como a desordem era frequentemente concebida e vivida - vivida [...] como a inversão dos valores hierárquicos, ou a perturbação dos processos normais da natureza, ou a proclamação de eventos apocalípticos, ou as consequências da idolatria e do pecado. Nessas categorizações, o demonismo já estava implícito e as bruxas não ficavam muito atrás.

¹⁸ Em um período de aproximadamente 100 anos, ocorreram os seguintes eventos: a transição do modo de produção feudal para o capitalismo (mercantilismo), a era das grandes navegações e a consequente instauração do comércio em escala mundial, a solidificação das monarquias nacionais, a reforma protestante e a contrarreforma católica.

(CLARK, 2006, p.693–694)

Vale observar que a preocupação do aldeão naquelas situações era dotada de um aspecto pragmático. Ele pouco compreendia as nuances da complexa doutrina da bruxaria. Seu desejo era simples: cessar com aquilo que lhe estava causando algum mal ou prejuízo — a morte de uma vaca, o mau tempo que prejudicava a sua colheita etc.

Quanto à falta de espírito crítico em relação a tal estado de repressão jurídico-teológica, poucas obras apresentavam argumentos que discordassem dos conceitos constituintes da doutrina da bruxaria. Uma das mais famosas foi o “De Praestigiis Daemonum” de Johann Weyer, publicado em 1563. Nessa obra, entre outros argumentos, Weyer advogava que a confissão dos *maleficia* pelas bruxas poderia ser explicada por causas naturais: tratava-se de uma doença — a melancolia — que estaria presente nos úteros das confessoras. No entanto, o autor jamais afastou as influências sobrenaturais e demoníacas sobre as acusadas: o diabo era o responsável por fazê-las acreditarem na própria culpa sobre os atos a elas imputados.

Jean Bodin, na obra “La Démonomanie des Sorciers”, em 1580, dedicou um anexo inteiro no combate aos argumentos de Weyer. Para ele, a conspiração diabólica era clara, uma vez que as confissões dos acusados, independentemente da região, eram todas idênticas. Além disso, o perigo descrito nessas confissões era tão manifesto que estávamos autorizados a utilizar todas as armas disponíveis (ameaças, torturas, mentiras, espionagem etc.) no sentido de combatê-las. Dessa forma, contra crimes tão antinaturais e extraordinários, Bodin (2001) advogava, sem meias-palavras, um claro procedimento jurídico de exceção.

Um divertido desafio ao argumento de que as confissões praticamente idênticas eram prova definitiva da existência dos *maleficia* foi apresentado por Friedrich Spee, na obra “Cautio Criminalis”, publicada em 1631. Ora, sob tortura, ninguém se furtaria a confessar o que o inquisidor desejasse ouvir, e nesse rol poderíamos incluir, entre outros, os membros do clero, juízes e, inclusive, conforme admite, o próprio Spee.

Igualmente, no ensaio “Dos Coxos”, encontramos uma forte crítica à doutrina da bruxaria. Cuidadosamente, Montaigne descreve no texto uma série de ocasiões em que a notória inconstância de testemunhos, seja em julgamentos ou nas mais prosaicas situações, afetava a correta avaliação do que se buscava esclarecer, conduzindo-nos invariavelmente ao erro. Na maioria das vezes, segundo o filósofo,

supostos eventos fantásticos poderiam ser explicados de modo simples e natural, sem a necessidade de se fazer qualquer concessão a concepções místicas ou extraordinárias. Entretanto, a doutrina da bruxaria produzia argumentos poderosos, minando qualquer pretensão de uma discussão mais equilibrada. Por exemplo, ao discutir a transformação de homens em animais ou em pedra por encantamentos de bruxas (dada a sua natureza, um evento de difícil credibilidade) Bodin argumentava:

E se admitirmos a verdade da história sagrada em Daniel [que relatou a transformação de Nabucodonosor em uma criatura semelhante a uma besta], que não pode ser colocada em dúvida, e da história da mulher de Ló transformada em pedra imóvel, é certo que a transformação do homem em boi ou pedra é possível, assim como em muitos outros animais. (BODIN, 2001, p.129)

O argumento estabelece uma paridade entre os eventos miraculosos relatados na Bíblia e os eventos extraordinários atribuídos à prática da bruxaria, especialmente no que diz respeito ao caráter fantástico e sobrenatural de ambos. Assim, se não questionamos a ocorrência dos primeiros, não deveríamos duvidar da existência dos últimos. No ensaio "Dos Coxos", Montaigne apresenta um contra-argumento contundente, ao chamar a nossa atenção para a relação entre o evento sobrenatural e o seu respectivo relato ou testemunho. No caso dos milagres bíblicos, o testemunho vinha daqueles escolhidos por Deus para falar em seu nome. Por seu turno, quando se tratava de eventos sobrenaturais atribuídos à feitiçaria, o testemunho era geralmente fornecido por pessoas comuns que se encontravam envolvidas no processo criminal. Embora não houvesse razão para duvidar da veracidade da conexão entre os milagres bíblicos e os testemunhos endossados por Deus, a conexão entre os eventos sobrenaturais relacionados à bruxaria e o caráter mundano do respectivo testemunho não ofereceria tal garantia. Portanto, conclui Montaigne, em relação aos relatos de supostos eventos relacionados à prática da bruxaria, deveríamos, ao encontrá-los, sempre os contrapor a causas simples e naturais ou, na sua ausência, suspender nosso julgamento.¹⁹

¹⁹ Um século e meio após a publicação do ensaio de Montaigne, David Hume propõe uma série de argumentos céticos em relação aos testemunhos de eventos miraculosos. Em sua argumentação, o filósofo retoma a ideia montaigniana da relação entre o evento miraculoso e o seu respectivo testemunho; no entanto, ele a estabelece agora sobre o pano de fundo da nossa experiência regular e consistente. Em seu livro "Investigações sobre o Entendimento Humano e sobre os Princípios da Moral", mais precisamente no capítulo intitulado "Dos Milagres", Hume (2004, p.161) afirma que "nenhum testemunho é suficiente para estabelecer um milagre, a menos que seja de um tipo tal que

De acordo com Koyré, argumentos como aqueles apresentados por Bodin e a sua matriz, a doutrina da bruxaria, seriam consequências da destruição da chamada “síntese aristotélica”. Fato que colaborava decididamente para que aquela doutrina dominasse sem muito esforço a cena discursiva em que se encontravam os membros da elite europeia naquele momento.

A grande inimiga da Renascença, do ponto de vista filosófico e científico, foi a síntese aristotélica, e pode dizer-se que sua grande obra foi a destruição dessa síntese.

Ora, esses traços que acabo de evocar, a credulidade, a crença na magia, etc., me parecem ser consequências diretas daquela destruição. Com efeito, depois de ter destruído a física, a metafísica e a ontologia aristotélicas, a Renascença se viu sem física e sem ontologia, isto é, sem possibilidade de decidir, de antemão, se alguma coisa é possível ou não.

Ora, parece-me que, no nosso pensamento, o possível sempre se sobrepõe ao real, e o real não é senão o resíduo desse possível. Ele se coloca ou se acha no contexto do que não é impossível. No mundo da ontologia aristotélica, há uma infinidade de coisas que não são possíveis, uma infinidade de coisas, portanto, que sabemos de antemão serem falsas.

Uma vez essa ontologia destruída, e antes que uma nova ontologia, elaborada somente no século XVII, seja estabelecida, não se dispõe de critério algum que permita decidir se a informação que se recebe de tal ou qual “fato” é verdadeira ou não. Daí resulta uma credulidade sem limites. (KOYRÉ, 1982, p. 47)

Com base no que desenvolvemos até agora, podemos inferir que especialmente os integrantes da elite europeia se encontravam, após o estabelecimento da doutrina da bruxaria, circunscritos a uma bolha epistêmica. Vale lembrar como a ampla circulação de livros proporcionada pelo mercado editorial naquela época e a intensa vida social, restrita apenas a seus membros, contribuíram decisivamente para a divulgação e amplificação de opiniões de caráter fantástico e bizarro. Tal fato colaborava para a consolidação da homogeneidade de pensamento,

sua falsidade seja ainda mais extraordinária do que o próprio fato que se propõe a estabelecer". Dessa forma, segundo o filósofo, se alguém afirmou ter visto um homem ressuscitar, devemos nos perguntar: o que seria mais extraordinário, o evento ou a falsidade do testemunho? Dado que observamos, segundo a nossa experiência regular e consistente, a frequente ocorrência de relatos falsos, não haveria nada de extraordinário na falsidade de um testemunho de alguém que afirmasse ter visto um homem ressuscitar. Assim, o testemunho desse indivíduo não seria suficiente para a comprovação do suposto milagre. Vale ressaltar que a argumentação de Hume não sugere que milagres não ocorram, mas sim que nenhum testemunho é suficiente para comprová-los.

que induzida pelo medo da perda de controle social, transmutava-se rapidamente em radicalização.

Vale destacar também como vozes discordantes eram ignoradas ou mesmo violentamente silenciadas. Em relação a este último aspecto, podemos constatar que a situação era extremamente delicada e até perigosa, tanto para Montaigne quanto para todos aqueles que discordavam da doutrina da bruxaria. Embora houvesse um argumento refinado e um ceticismo sofisticado em seu ensaio, faltavam crenças estabelecidas dentro de seu grupo social que permitissem ao filósofo desafiar e combater aqueles discursos falaciosos com a energia necessária. Em situações de evidente tensão discursiva, é essencial ter à disposição um conjunto rival de teorias que possam sustentar argumentos ou explicações alternativas. Na ausência de tal conjunto, ou se este for pouco conhecido ou pouco desenvolvido, corre-se o risco de transformar qualquer discordância em um autêntico *tour de force* discursivo. Era exatamente isso que acontecia durante o reinado da doutrina da bruxaria e da consequente bolha epistêmica que se formou naquela época. Não havia uma visão mecanicista sólida para sustentar explicações naturais e simples para eventos supostamente fantásticos. Qualquer conjunto de argumentos mais equilibrado que contrariasse as crenças compartilhadas por aquele grupo era desconsiderado ou sufocado.

Somente a partir do Iluminismo, quando o pensamento mecanicista se impôs sobre a elite culta, instituindo novas formas de compreensão dos fenômenos naturais e um novo papel para Deus nessa nova disposição, é que as teses da doutrina da bruxaria se tornaram expressão da mais pura superstição. O que permitiria a Voltaire (2011) ironicamente se mostrar saudosos da época em que eventos fantásticos eram vistos por toda parte; e de apontar, de modo mordaz, na frivolidade e crueldade da nobreza, uma explicação para a ocorrência e a sustentação de crenças tão sombrias e trágicas — elas proporcionavam um solo fértil para o surgimento de tantas histórias fantásticas e bizarras que divertiam e exercitavam a mente dos integrantes da elite.

3.2 Winston Parva

Ao final da década de 1950, os sociólogos John Scotson e Norbert Elias iniciaram uma pesquisa de campo em uma pequena comunidade próxima à cidade de Leicester na Inglaterra. O objetivo da pesquisa era o estudo do crescimento da delinquência juvenil na região. No decorrer das investigações, a pesquisa acabou tomando outro rumo, transformando-se em um vigoroso estudo sobre as relações de poder entre um grupo de moradores que residia há muito tempo na região, “os estabelecidos”, e outro de residência recente, “os *outsiders*”. Em 1965, os autores divulgaram o resultado das investigações, publicando a obra “Os Estabelecidos e os *Outsiders*”.

Ao começarmos a investigar os fatos e buscar explicações, nosso interesse deslocou-se dos diferenciais de delinquência para as diferenças de caráter desses bairros e para as relações entre eles. No decorrer de uma exploração bastante minuciosa do microcosmo de Winston Parva, com seus três bairros distintos, passamos a conhecer bastante bem a localidade e alguns de seus membros. O fascínio exercido em nós por seus problemas foi aumentando sistematicamente — sobretudo à medida que percebemos, pouco a pouco, que alguns deles tinham um caráter paradigmático: lançavam luz sobre problemas comumente encontrados, em escala muito maior, na sociedade como um todo. (ELIAS; SCOTSON, 2000, p.15)

A comunidade de Winston Parva, nome fictício, dividia-se em três bairros. Dois deles eram compostos por pessoas que residiam há duas ou três gerações no local, e outro, por operários que moravam ali há pouco tempo. O que chamava a atenção dos autores era a divisão entre os membros dos dois grupos. Ou seja, os moradores que residiam na área de ocupação mais antiga, que se autodenominavam “moradores da aldeia”, discriminavam e estigmatizavam os moradores de ocupação mais recente, denominados pejorativamente de “moradores do loteamento” por aqueles. Para os “aldeões”, os “moradores do loteamento” seriam destituídos de qualquer valor moral superior, não possuindo, por esse motivo, o necessário carisma grupal distintivo. A partir dessa divisão em grupos, descobrir a origem do sentimento de superioridade dos moradores mais antigos, bem como determinar quais seriam os meios que dispunham para dominar os mais novos, passou a ser o objetivo da pesquisa.

Naquela pequena comunidade, a superioridade de forças do grupo estabelecido desde longa data era desse tipo, em grande medida. Baseava-se no alto grau de coesão de famílias que se conheciam havia duas ou três gerações, em contraste com os recém-chegados, que eram estranhos não apenas para os antigos residentes como também entre si. Era graças a seu maior potencial de coesão, assim como à ativação deste pelo controle social, que os antigos residentes conseguiam reservar para as pessoas de seu tipo os cargos importantes das organizações locais, como o conselho, a escola ou o clube, e deles excluir firmemente os moradores da outra área, aos quais, como grupo, faltava coesão. (ELIAS; SCOTSON, 2000, p.22)

Os estabelecidos detinham os postos de poder e prestígio na região, reafirmando com isso seu carisma grupal. Os *outsiders*, devido à sua pouca coesão, não conseguiam se organizar nem se impor enquanto grupo. Dominados, eram frequentemente estigmatizados pelos estabelecidos. Inclusive, com o decorrer do tempo, passaram a internalizar tal marca, acreditando-se realmente inferiores aos aldeões.

Os mais “antigos” muitas vezes conseguem impor aos recém-chegados a crença de que estes são inferiores ao grupo estabelecido, não apenas em termos de poder, mas também “por natureza”. E essa internalização da crença depreciativa do grupo socialmente superior pelo socialmente inferior, como parte da consciência e da imagem que este tem de si, reforça vigorosamente a superioridade e a dominação do grupo estabelecido. (ELIAS; SCOTSON, 2000, p.175)

Além da antiguidade no local, não se observava fatores econômicos, étnicos, religiosos ou educacionais que justificassem a coesão grupal dos estabelecidos e a estigmatização dos *outsiders*:

Andando pelas ruas das duas partes de Winston Parva, o visitante ocasional talvez se surpreendesse ao saber que os habitantes de uma delas julgavam-se imensamente superiores aos da outra. No que concerne aos padrões habitacionais, as diferenças entre as duas áreas não eram particularmente evidentes. Mesmo examinando essa questão mais de perto, era surpreendente, a princípio, que os moradores de uma área tivessem a necessidade e a possibilidade de tratar os da outra como inferiores a eles e, até certo ponto, conseguissem fazê-los sentirem-se inferiores. Não havia diferenças de nacionalidade, ascendência étnica, “cor” ou “raça” entre os residentes das duas áreas, e eles tampouco diferiam quanto a seu tipo de ocupação, sua renda e seu nível educacional — em suma, quanto a

sua classe social. As duas eram áreas de trabalhadores. (ELIAS; SCOTSON, 2000, p.21)

Ao serem incapazes de enquadrar aquela situação factual segundo os referenciais teóricos que normalmente balizam objetos dessa natureza, os autores se viram compelidos a explorar novos caminhos e novas formulações. Esta oportunidade única de pesquisa possibilitou, assim, a proposta de um novo paradigma empírico de relacionamento entre grupos: a figuração estabelecidos-*outsiders*.

Ali se encontrava, sob forma particularmente pura, uma fonte de diferenciais de poder entre grupos inter-relacionados que também desempenha um papel em muitos outros contextos sociais, mas que, nestes, muitas vezes se dissimula, aos olhos do observador, por outras características marcantes dos grupos em questão, tais como a cor ou a classe social. Mediante um exame mais detido, é frequente poder-se descobrir que, também nesses outros casos, tal como em Winston Parva, um grupo tem um índice de coesão mais alto do que o outro e essa integração diferencial contribui substancialmente para seu excedente de poder; sua maior coesão permite que esse grupo reserve para seus membros as posições sociais com potencial de poder mais elevado e de outro tipo, o que vem reforçar sua coesão, e excluir dessas posições os membros dos outros grupos — o que constitui, essencialmente, o que se pretende dizer ao falar de uma figuração estabelecidos-*outsiders*. (ELIAS; SCOTSON, 2000, p.22)

A figuração concentra-se na ideia de que uma maior coesão grupal estaria na raiz das inúmeras situações em que se observa alguma forma de dominação entre grupos. Uma maior coesão permite a imposição de um grupo sobre outro, possibilitando ao mais coeso um acesso mais amplo às instâncias disponíveis de poder e opressão. O que, por seu turno, reforça a coesão grupal já existente, demandando conseqüentemente mais poder.

Vale notar que de modo surpreendente, em Winston Parva, a fofoca permeava praticamente todas as relações de poder entre os residentes. Ela reforçava, em sua versão elogiosa, a coesão grupal dos estabelecidos, assim como em sua versão depreciativa, a estigmatização dos *outsiders*.

[...] o grupo estabelecido atribuía a seus membros características humanas superiores; excluía todos os membros do outro grupo do contato social não profissional com seus próprios membros; e o tabu em torno desses contatos era mantido através de meios de controle social como a **fofoca elogiosa** [praise gossip], no caso dos que o observavam, e a ameaça de **fofocas depreciativas** [blame gossip]

contra os suspeitos de transgressão. (ELIAS; SCOTSON, 2000, p.20, **negrito nosso**)

A fofoca, em outras palavras, não é um fenômeno independente. O que é digno dele depende das normas e crenças coletivas e das relações comunitárias. A imagem negativa do loteamento [os *outsiders*], que levava os “aldeões” [os estabelecidos] a perceberem como um mexerico bem-vindo qualquer incidente que servisse de confirmação dela, era o inverso da imagem positiva que eles tinham de si mesmos. (ELIAS; SCOTSON, 2000, p.121)

A fofoca sustentava-se, no caso de Winston Parva, nas crenças e nas atitudes compartilhadas por aqueles que a disseminavam — os estabelecidos. Vale observar que, como consequência do processo de estigmatização, a fofoca depreciativa, cujo principal alvo eram os *outsiders*, era aceita por estes como a expressão da verdade em muitas situações, como mencionamos.

3.2.1 A fofoca: sua função social e epistêmica

Enquanto fator de reforço da coesão grupal, como algo tão prosaico e até moralmente discutível pode ser tão importante? Coady (2006) analisa este, além de outros aspectos relacionados à fofoca. Para o autor, ela apresenta as seguintes características:

- a) aqueles que a transmitem acreditam na sua verdade;
- b) sua transmissão sustenta-se em crenças perfeitamente justificáveis;
- c) não necessita ser maliciosa, embora às vezes seja;
- d) seu conteúdo é invariavelmente pessoal, embora possa se referir a pessoas que representem ou tipifiquem determinados grupos.

Ademais, segundo Coady (2006, p.255, tradução nossa), "a fofoca não exhibe nenhuma diferença epistêmica essencial em relação ao testemunho comum". O autor assevera ainda que as possíveis críticas ao valor epistêmico da fofoca se concentram no fato de que o fofocueiro busca muitas vezes apresentar-se como detentor de informações privilegiadas, ou como alguém próximo a uma fonte confiável de poder que possa acessá-las. Entretanto, tais motivações também animam aqueles que testemunham nos tribunais:

Testemunhas que compartilham informações sem indícios de distorção podem, no entanto, ser motivadas pelo desejo de se destacarem em grandes eventos, deleitando-se com a emoção que seus relatos causarão. Elas também podem ser impulsionadas por um compromisso de acertar as coisas - qualquer relato que inspire um testemunho confiável se apoia na existência de motivações mistas. Em segundo lugar, mesmo que as motivações associadas à fofoca não sejam das mais nobres, elas não precisam descartar a preocupação com a verdade. O desejo de impressionar pode, de fato, tornar a pessoa mais escrupulosa ao transmitir informações precisas, uma vez que o status de informante confiável pode ser crucial para a apreciação positiva que se busca. (COADY, 2006, p.261, tradução nossa)

Outra crítica feita ao valor epistêmico da fofoca é que, devido à sua própria natureza, não haveria nenhum modo de testar a sua confiabilidade. Ou seja, não existiria nenhuma metodologia disponível que pudesse confirmá-la, falsificá-la ou corrigi-la. No entanto, de acordo com Coady, quando o fofocueiro é próximo da vítima da fofoca e não possui nenhum interesse em mentir, isso já seria suficiente para garantir a confiabilidade do conteúdo fofocado. Nesses casos, a confiabilidade estaria assegurada independentemente da vítima, uma vez que, em muitas situações, ela não tem interesse, possibilidade ou mesmo condição de confirmar ou refutar o conteúdo fofocado – seja porque não deseja oferecer nenhuma declaração em relação ao assunto, seja porque suas ações não foram a origem da fofoca.

Rosnow (2001) destaca, por seu turno, uma outra função da fofoca: o reforço das normas sociais.

[...] quer as pessoas sejam motivadas por uma conformidade intencional ou sutil com o grupo, ou com a comunidade, ou com a cultura, as consequências do boato e da fofoca também refletem até que ponto as normas sociais podem ser impostas a indivíduos que ostensivamente ameaçam ou violam-nas. Em pequenos grupos, por exemplo, a fofoca pode ser uma forma de guiar o rebanho, dizendo: "Estes são os limites, e você está ultrapassando-os". (ROSNOW, 2001, p.224, tradução nossa)

Nesse aspecto, Coady acrescenta que a fofoca entendida como controle, acaba servindo muitas vezes de veículo à discriminação e à perseguição de minorias, como acontecia em Winston Parva. O autor também observa que em outras situações, longe de ser um instrumento de controle, a fofoca é empregada como um elemento de transgressão social. Esse fato foi observado, por exemplo, na França pré-

revolucionária do século XVIII, onde a rígida estrutura de controle foi gradualmente minada pela fofoca.²⁰

Em outra perspectiva interpretativa, ao considerar os aspectos epistêmicos envolvidos, Bertolotti e Magnani (2014) afirmam que a fofoca apresentar-se-ia também como um processo colaborativo grupal:

[...] a fofoca diz respeito, em última análise, à atualização e ao compartilhamento da base de conhecimento coletivo de um grupo. Base, por sua vez, composta de informações compartilhadas sobre as pessoas, bem como da avaliação comum (também moral) dessas mesmas informações. (BERTOLOTTI; MAGNANI, 2014, p.4045, tradução nossa)

Dessa forma, manter um grupo unido seria impossível sem um processo altamente eficiente de troca de informações. Processo muitas vezes impulsionado e sustentado pela fofoca ou, de modo mais preciso, por inferências de natureza moral. Neste último aspecto, ainda segundo Bertolotti e Magnani, a forma inferencial da fofoca apresentar-se-ia conforme a estrutura de um **raciocínio abduutivo** ou **hipótese explicativa provisória**, nos termos propostos por Peirce (1974).²¹

²⁰ Em uma sociedade em que os meios de comunicação formais eram estereotipados e rigidamente controlados pelos detentores do poder, a principal fonte de informação sobre o rei e as elites dominantes vinha das fofocas e dos rumores [...]. [As] fofocas e rumores sobre a Corte Francesa não apenas forneciam uma rara fonte de informações pessoais e políticas, mas também, possivelmente, desempenharam um papel na preparação do terreno para a Revolução. (COADY, 2006, p.258, tradução nossa)

²¹ "Muito antes de eu ter classificado a abdução como inferência, já os lógicos tinham reconhecido que a operação de adotar uma hipótese explicativa estava sujeita a certas condições. A saber, a hipótese não pode ser admitida, mesmo como hipótese, a não ser que dê conta dos fatos, pelo menos alguns deles. A forma da inferência, portanto, é esta: O fato surpreendente C é observado; Ora, se A fosse verdade, C seria um fato natural. Assim, há razão para suspeitar que A é verdadeiro. Assim, A não pode ser abduktivamente inferido, ou, se preferirem, abduktivamente conjecturado até que o seu total conteúdo se ache presente na premissa "Se A fosse verdade, C seria um fato natural". (PEIRCE, 1974, p.60)

Exemplificando. Numa manhã de segunda-feira, observamos o seguinte diálogo no escritório de um banco de investimentos:

(Pedro)

(1) — Parece que o Manuel foi finalmente demitido.

(João)

(2) — Nossa! Isso seria ótimo! Você tem certeza?

(Pedro)

(3) — Só pode ter ocorrido! Na última quarta-feira, você se lembra? A diretoria nos informou que estava avaliando as recentes reclamações de assédio moral encaminhadas à ouvidoria, e caso fossem confirmadas as acusações, sérias providências seriam tomadas. Desde aquele informe, ninguém mais viu o Manuel. E nesse exato instante, o pessoal da área que ele coordena está em reunião com o diretor, o que é bastante incomum para uma segunda-feira!

Dessa forma, concluem os autores, a fofoca compreendida como um raciocínio abduativo apresentaria um duplo aspecto:

No que diz respeito ao seu valor positivo, definir a fofoca como uma atividade abduativa enfatiza sua dimensão epistemologicamente criativa: é sensato afirmar que a fofoca é uma atividade cognitiva profundamente enraizada na ecologia social de alguém [...]. No que diz respeito ao seu valor negativo, rotular a fofoca como uma atividade abduativa explicita o caráter constitutivamente experimental e provisório do conhecimento que ela fornece. (BERTOLOTTI; MAGNANI, 2014, p.4055–4056, tradução nossa)

De qualquer modo, como vemos, a fofoca apresenta um caráter multifacetado. De um ponto de vista sociológico, poderia ser entendida como um reforço dos controles sociais de uma comunidade ou mesmo como um elemento transgressor desses mesmos controles em outras situações. Por outro lado, naquilo que mais nos interessa, de um ponto de vista epistêmico, a fofoca poderia ser entendida como um testemunho ou ainda como uma inferência de cunho abduativo. Nesse último sentido, enquanto **hipótese explicativa provisória**, ao fluir no interior de um grupo, alimentaria a base de conhecimento coletivo.

Além de ter inspirado a figuração estabelecidos-*outsiders*, a comunidade de Winston Parva possuía, nos termos discutidos em nosso estudo, os distintivos de uma bolha epistêmica. Vale observar que Elias e Scotson elencam inclusive as possíveis razões para tal radicalização:

Com ou sem razão, eles, como muitos outros grupos estabelecidos, sentiram-se expostos a um ataque tríplice — contra seu monopólio das fontes de poder, contra seu carisma coletivo e contra suas normas grupais. Repeliram o que vivenciaram como um ataque, cerrando suas fileiras contra os recém-chegados, excluindo-os e humilhando-os. (ELIAS; SCOTSON, 2000, p.50)

Assim, segundo o raciocínio abduativo: O fato surpreendente C é observado: o sumiço do Manuel e a reunião dos funcionários da sua área com o diretor na segunda-feira.

Se A fosse verdade, C seria um fato natural: se o Manuel tivesse sido demitido, os fatos relatados pelo proferimento (3) seriam naturais. Assim, há razão para suspeitar que A seja verdadeiro, isto é, a demissão de Manoel deve ter realmente ocorrido.

Antes de se unirem contra os *outsiders*, dada a sua antiguidade na região, os estabelecidos já se apresentavam como um grupo bastante coeso cujos membros pensavam de modo semelhante. A polarização, tanto das crenças compartilhadas, quanto do consequente discurso explicativo que se apresentava configurado como hipótese explicativa provisória (fofoca), começou a se formar a partir da chegada dos *outsiders*.

Vale ainda observar outro aspecto: se a fofoca era invariavelmente utilizada na estigmatização dos *outsiders*; no entanto, em alguns momentos, também se direcionava contra os indivíduos não-conformistas que residiam entre os “estabelecidos”:

Foram encontrados, vez por outra, durante as entrevistas, exemplos dessa exclusão de pessoas que a opinião pública da “aldeia” [os estabelecidos] suspeitava de “não conformismo” embora, de modo geral, a “gente boa”, sobretudo a “gente boa” da área de elite, procurasse ocultar a presença de “ovelhas negras” sociais em sua rua. Assim é que, ao final de uma entrevista na área de “elite” da “aldeia”, uma dona-de-casa de meia-idade perguntou se poderia saber quem seria o próximo entrevistado. Quando informada, disse: “Oh! Eu não iria lá, não! Vá ao número 15, eles são gente boa, mas lá, não, ela é uma leviana, só está aqui há um ano. Vá à casa dos Sewell, eles são boa gente”. (ELIAS; SCOTSON, 2000, p.81)

Qual seria o motivo da próxima entrevistada não ser considerada "gente boa"?

[Segundo a própria entrevistada, alvo da fofoca:] “Eles são muito reservados. Falam comigo na rua, mas é só.” Em seguida, contou que havia convidado “o lixeiro para uma xícara de chá num dia frio”, logo depois de se mudar para Winston Parva. “Eles viram. Isso chocou o pessoal daqui.” Não era apenas que os recém-chegados tivessem que adotar os padrões da “aldeia”: eles também tinham que mostrar ostensivamente que os adotavam. Caso contrário, eram colocados num nível inferior na hierarquia de status das famílias “aldeãs” e tratados como *outsiders*. (ELIAS; SCOTSON, 2000, p.82)

A fala da senhora indica o constrangimento social a que eram submetidos os moradores da aldeia que não se comportavam de acordo com o padrão estabelecido. Mais precisamente, os instrumentos de constrangimento seriam a fofoca depreciativa e a exclusão social.

[...] é evidente que os “aldeões” formavam, em muito maior grau, um grupo relativamente fechado. Tinham desenvolvido tradições e padrões próprios. Quem não cumpria essas normas era excluído como sendo de qualidade inferior.

[...] Acima de tudo, desenvolveram como arma uma “ideologia”, um sistema de atitudes e crenças que enfatizava e justificava sua própria superioridade, e que rotulava as pessoas do loteamento como sendo de categoria inferior. Construída em torno de alguns temas estereotipados, sua ideologia de status disseminou-se e foi mantida por um fluxo constante de fofocas, que se agarrava a qualquer acontecimento da “aldeia” que pudesse ajudar a engrandecer a comunidade “aldeã” e a qualquer acontecimento, entre as pessoas da outra zona, capaz de reforçar a imagem negativa do loteamento. Ela também ajudou a bloquear a percepção de qualquer acontecimento que tivesse alguma possibilidade de contradizê-la. (ELIAS; SCOTSON, 2000, p.65)

O reforço da visão grupal por meio da disseminação da fofoca, a delimitação da atuação dos membros aos limites da aldeia, o sufocamento das opiniões e das atitudes discordantes ao comportamento padrão por meio do isolamento social e da fofoca depreciativa, esses são aspectos que, segundo o nosso ponto de vista, fazem do cenário estudado em Winston Parva um bom exemplo de uma bolha epistêmica.

3.3 As mídias sociais

A campanha do plebiscito que resultou na saída do Reino Unido da União Europeia (BREXIT) em 2016; as campanhas presidenciais americanas e brasileiras respectivamente em 2016 e em 2018; a criminosa difusão de toda a sorte de teorias pseudocientíficas durante o auge da pandemia de Covid-19. O que esses eventos teriam em comum? Resposta: uma intensa polarização grupal nas redes sociais alimentada por muita desinformação e ódio.

Disputa-se ainda se as causas que propiciaram os desdobramentos trágicos desses acontecimentos originaram-se na radicalização virtual. Contudo, pode-se afirmar que sem esse tipo de estrutura que o atual ambiente digital apresenta, tais eventos dificilmente teriam o alcance e a influência deletéria que exibiram.

Desde o seu estabelecimento, baseado na exploração do potencial oferecido pela internet no início do século XXI, diversos eventos significativos colaboraram para a formação do mercado digital como atualmente conhecemos. A primeira grande mudança expressiva ocorreu por volta do ano de 2003, quando a empresa de tecnologia Google introduziu uma alteração em seu sistema de buscas de informações

na internet, revolucionando o modelo de negócios das empresas que atuavam no setor. Antes dessa alteração, os dados referentes às pesquisas dos usuários, coletados a cada interação do buscador de informações da Google, eram utilizados apenas para aprimorar a própria ferramenta. No entanto, a partir daquele momento, a empresa passou a coletar também informações pessoais de seus usuários, como padrões de pesquisa, localização, dados pessoais, entre outras. Com o uso desse excedente informacional, desse **superávit comportamental**, tornou-se possível a categorização dos usuários em grupos, tanto em relação as suas características pessoais quanto ao seu comportamento. A ideia era simples: ao segmentar seus usuários, a empresa conseguiria direcionar propagandas de acordo com as características e afinidades de cada grupo. De acordo com Zuboff (2021), este foi o primeiro passo na instauração de um novo modelo de negócios, ou melhor, de um novo paradigma de produção econômica: o **capitalismo de vigilância**. Vale ressaltar que não haveria nenhum processo de categorização nem capitalismo de vigilância sem o uso daquele *superávit* comportamental coletado gratuitamente a partir do comportamento dos usuários.

O capitalismo de vigilância reivindica de maneira unilateral a experiência humana como matéria-prima gratuita para a tradução em dados comportamentais. Embora alguns desses dados sejam aplicados para o aprimoramento de produtos e serviços, o restante é declarado como *superávit* comportamental do proprietário, alimentando avançados processos de fabricação conhecidos como “inteligência de máquina” e manufaturado em produtos de predição que antecipam o que um determinado indivíduo faria agora, daqui a pouco e mais tarde. (ZUBOFF, 2021, p.23)

À medida que novas tecnologias de rede e poderosos algoritmos foram desenvolvidos, as empresas de tecnologia aprimoraram a categorização de seus usuários, gerando e mantendo em sua posse perfis comportamentais cada vez mais detalhados. De um modo inédito, dada a grande quantidade de informações coletadas, as empresas, por meio de sofisticadas ferramentas de previsão comportamental, podiam agora comercializar a conduta futura de seus usuários.

Por fim, esses produtos de predições são comercializados num novo tipo de mercado para predições comportamentais que chamo de *mercados de comportamentos futuros*. Os capitalistas de vigilância têm acumulado uma riqueza enorme a partir dessas operações

comerciais, uma vez que muitas companhias estão ávidas para apostar no nosso comportamento futuro. (ZUBOFF, 2021, p. 23–24)

Após a implantação desse novo modelo, a empresa de tecnologia Google, que em 2004 movimentava um milhão de dólares por dia, aumentou esse valor em trinta vezes até 2010. Como se esperava, diante desse sucesso, outras empresas passaram a adotar o mesmo modelo em seus serviços virtuais. Entre estes, começava a se destacar o *Facebook*. No entanto, este serviço não se notabilizava por buscar informações na internet. Nos primeiros anos do século XXI, o *Facebook* era o representante de um novo tipo de serviço que começava a dominar o mercado virtual: os Serviços de Redes Sociais, doravante, SRS.

A procura por maneiras de se comunicar por meio da internet sempre fora uma preocupação. Um dos desenvolvimentos mais exitosos nesse sentido, sendo inclusive bastante utilizado até hoje, foi o correio eletrônico, popularmente conhecido como e-mail. E o seu sucesso foi tão grande que praticamente toda a comunicação interpessoal virtual se dava por meio desse serviço no final do século passado. Segundo previsões naquela época, o próximo passo nas comunicações interpessoais seria no sentido da fuga do real, ou seja, os relacionamentos sociais ocorreriam em mundos virtuais e entre indivíduos fictícios:

Outro equívoco sobre “comunidades virtuais” comum na década de 1990 era que a atividade on-line era algo separado da vida “real”. Sem os marcadores usuais de presença física, a identidade online parecia estar separada do cotidiano, quase infinitamente maleável. Fazer essa divisão binária entre a vida on-line e off-line tende a incentivar a banalização ou a exaltação das atividades on-line. Ficar online torna-se uma fuga covarde da vida “real” ou uma entrada em êxtase em um novo mundo de infinitas possibilidades. (CHESHER, 2020, p.108, tradução nossa)

No entanto, a fuga do real não ocorreu, pelo menos não na dimensão que se esperava:

Essas suposições iniciais de que as pessoas seriam atraídas a brincar com sua identidade foram contraditadas pela vigorosa adoção dos SRS nos anos 2000, onde a identidade “real” nos espaços on-line era quase sempre colocada em primeiro plano. A maioria das pessoas preferiu apresentar on-line o que percebem ser sua própria identidade

autêntica em vez de experimentar novas *personas*. Em vez de um jogo separado e sobrenatural, o uso de mídia de rede se tornaria outro aspecto da vida cotidiana. (CHESHER, 2020, p.108–109, tradução nossa)

A partir do momento em que o próprio usuário começou a compartilhar de modo consensual a sua privacidade e até mesmo a sua intimidade, o *superávit* comportamental, de mera informação residual, ascendeu a outro nível, transformando-se em uma *commodity* de alto valor. O que acabaria por direcionar um grande volume de investimentos no estudo e no desenvolvimento de novas e sofisticadas ferramentas de mineração e exploração dessas informações. Exemplos notórios desses desenvolvimentos foram a criação de duas novas funcionalidades no *Facebook*, a saber: o *feed* de notícias e o botão “curtir” (*like*). Segundo informações da empresa Meta, proprietária do *Facebook*:

O *Feed* de Notícias de cada pessoa é uma coleção personalizada e em constante mudança de fotos, vídeos, links e atualizações de amigos, familiares, grupos, empresas e criadores de conteúdo aos quais elas se conectaram no *Facebook*.²²

Ademais, segundo Messingschlager e Holtz (2020), as informações disponíveis no *feed* de notícias de um usuário são trabalhadas pelo *Facebook*, passando necessariamente por um processo sucessivo de filtragens. Tal processo é composto pelas seguintes etapas:

- primeira etapa — informações oriundas do círculo de contatos do usuário (amigos, páginas seguidas etc.) são separadas. Informações fora desse círculo são excluídas;
- segunda etapa — além das mensagens selecionadas de acordo com o perfil do usuário na primeira etapa, um algoritmo curador, segundo sua lógica interna, automaticamente acrescenta novas informações relevantes, que são exibidas em primeiro lugar no *feed*;

²² META. **Meta Blueprint**, 2022. *Feed* de Notícias do *Facebook*: uma introdução para os criadores de conteúdo. Disponível em: [Facebook News Feed: An Introduction for Content Creators | Meta for Business](#). Acesso em: 29 dez. 2022.

- terceira etapa — as informações provenientes da segunda etapa são agora passíveis de serem selecionadas pelo usuário. E ao escolhê-las, o usuário atualiza seu perfil.

Desde a incorporação do *feed* de notícias ao *Facebook* em 2006, esse recurso se transformou no centro em torno do qual gravitaram praticamente todos os desenvolvimentos posteriores de novas funcionalidades. Esse é o caso do botão de “curtir”, que, desde a sua introdução em 2009, tem se mostrado uma ferramenta poderosa na exploração do *superávit* comportamental, aprimorando o conteúdo do *feed*:

A chefia do *Facebook* parece ter percebido apenas de forma gradual que o botão podia transformar a plataforma de livro em uma tempestade de espelhos, uma leitura passiva de uma enxurrada ativa de mútuos reflexos capazes de grudar os usuários a seus *Feeds* de Notícia. Assim, o botão *Curtir* era um espelho unidirecional do tamanho de um planeta, capaz de aumentar de maneira exponencial os suprimentos de matéria-prima. Quanto mais a usuária “curtia”, mais ela informava ao *Facebook* acerca do formato e da composição exatos de sua “mão”, permitindo com isso que a empresa ajustasse a luva continuamente, aumentando o valor preditivo dos sinais da usuária. (ZUBOFF, 2019, p.723)

Zuboff alerta-nos para esta importante característica algorítmica: à medida que o usuário explora e de alguma forma indica quais informações em seu *feed* o agradam (e, nesse aspecto, o botão de “curtir” potencializa essa ação), mais seu perfil comportamental é atualizado e moldado, calibrando as próximas ações do algoritmo:

A seleção de informações pelos usuários não determina apenas quais conteúdos eles leem no curto prazo. Em vez disso, os algoritmos utilizam as informações sobre essas decisões para personalizar ainda mais as informações apresentadas na próxima vez. As informações exibidas tornam-se cada vez mais consistentes com as próprias preferências e opiniões do usuário, enquanto informações que contradizem essas opiniões são exibidas apenas posteriormente ou até mesmo ocultadas. [...] Por exemplo, se uma pessoa clicou recentemente em alguns *links* com vídeos de cachorros, na próxima vez que visitar a rede social, ainda mais postagens sobre o assunto “cachorros” serão exibidas no topo da página. Por outro lado, essa pessoa provavelmente só verá vídeos de gatos se estiver procurando especificamente por eles. (MESSINGSCHLAGER; HOLTZ, 2020, p.95, tradução nossa)

Existe, portanto, um caráter recursivo entre a atualização do perfil comportamental do usuário e a disposição de informações no seu *feed*. Ou seja, escolhas pessoais moldam aquele perfil, sensibilizando o algoritmo curador a exibir novas informações afins em uma próxima vez. Estas, uma vez escolhidas, atualizam novamente o perfil, sensibilizando o algoritmo em um ciclo incessante. Pariser (2012) afirma que tal caráter algorítmico nos aprisiona no interior de autênticas bolhas-filtro:

O código básico no seio da nova internet é bastante simples. A nova geração de filtros on-line examina aquilo de que aparentemente gostamos – as coisas que fazemos, ou as coisas das quais as pessoas parecidas conosco gostam – e tenta fazer extrapolações. São mecanismos de previsão que criam e refinam constantemente uma teoria sobre quem somos e sobre o que vamos fazer ou desejar a seguir. Juntos, esses mecanismos criam um universo de informações exclusivo para cada um de nós – o que passei a chamar de bolha dos filtros – que altera fundamentalmente o modo como nos deparamos com ideias e informações. (Pariser, 2012, p.12)

As informações finais escolhidas por um usuário no seu *feed* de notícias provêm, como vemos, de duas fontes distintas: ou são fornecidas pelos seus próprios contatos, ou são induzidas pelo algoritmo curador. Quanto a esse último aspecto, um robusto estudo realizado por Bakshy *et al.* (2015) concluiu que, no caso do *Facebook*, a ação indutora do algoritmo curador não é tão significativa quanto se pensava. Ou seja, o modo de ser do usuário, as suas escolhas e crenças, bem como o perfil dos seus contatos, mostram-se mais decisivos na configuração final das informações consumidas. Este estudo envolveu uma amostra de aproximadamente dez milhões de pessoas que foram categorizadas, de acordo com suas preferências ideológicas, em dois grupos distintos: liberais e conservadores.

A metodologia do estudo consistiu em medir, durante o processo de filtragem, o fluxo de informações contrárias à tendência do perfil do usuário. Como mencionamos anteriormente, a despeito de concordarem ou não com o modo de pensar do usuário, sabemos que todas as informações geradas pelos seus contatos são reunidas em uma primeira etapa de filtragem. Após esse passo, as informações compiladas ficam à disposição da segunda etapa. Uma vez que em tese a ação curadora se processa neste ponto, caso pudéssemos medir a redução porcentual de informações discordantes, conseguiríamos avaliar a força da atuação algorítmica com precisão. Por exemplo, uma redução significativa na porcentagem de informações

discordantes nesta etapa resultaria em um *feed* de notícias fortemente influenciado pela curadoria algorítmica.

Inicialmente, os autores levantaram a quantidade total no *Facebook* das informações que de alguma forma contrariavam o perfil inicial dos usuários categorizados no estudo. De acordo com Bakshy *et al.*, considerando o total de informações que circulam nessa rede social sem mediação, cerca de 45% são informações discordantes em relação à orientação ideológica dos liberais, enquanto 40% das informações divergem dos posicionamentos apresentados pelos conservadores. Tendo em vista esses dados, conforme o estudo, após a primeira etapa de filtragem, os liberais passaram a compartilhar uma taxa de 24% de informações discordantes provenientes de seus contatos, enquanto os conservadores, uma taxa de 35%. E o que aconteceu após a segunda etapa, em que supostamente haveria uma ação mais efetiva da curadoria algorítmica?

Descobrimos que, após a classificação, há, em média, um pouco menos de conteúdo transversal [informações discordantes]: os conservadores veem aproximadamente 5% menos conteúdo transversal em comparação com o que os amigos compartilham, enquanto os liberais veem cerca de 8% menos conteúdo ideologicamente diverso. (BAKSHY *et al.*, 2015, p.1131, tradução nossa)

Finalmente, após a terceira etapa de filtragem:

O decréscimo estimado na probabilidade de um indivíduo clicar em um artigo transversal (informação discordante), uma vez que essa informação é exibida no *Feed* de Notícias, passa a ser de 17% para os conservadores e de 6% para os liberais. (BAKSHY *et al.*, 2015, p.1131, tradução nossa)

Em suma, na primeira etapa de filtragem, a porcentagem de informações discordantes diminuiu em 12,5% para os conservadores e em 46,67% para os liberais. Na segunda etapa, diminuiu em 5% para os conservadores e em 8% para os liberais. Na terceira etapa, a redução foi de 17% para os conservadores e de 6% para os liberais.

O estudo revelou, assim, que a porcentagem de informações discordantes em relação à orientação do usuário sofre um maior decréscimo na terceira etapa para os conservadores e na primeira para os liberais. Desse modo, a atuação do algoritmo

curador, que ocorre apenas na segunda etapa de filtragem, não é a principal responsável por uma maior homogeneidade de informações no *feed* do usuário. Mais precisamente, vale notar que essa atuação algorítmica em termos percentuais, independentemente do grupo de pertencimento do usuário, é bem menor que a soma das porcentagens das duas outras etapas. Portanto, o modo de ser e pensar do usuário e de seus contatos, bem como as crenças e as atitudes compartilhadas entre esses agentes, constituem as causas mais determinantes da homogeneidade no *feed* de notícias.

O estudo acima visava avaliar a influência do algoritmo curador no fluxo de informações discordantes em relação ao modo de pensar do usuário no *Facebook*. Em um estudo recente, sem se concentrar em uma mídia social específica, Min (2019) também se dispõe a avaliar o impacto dos algoritmos curadores, direcionando, no entanto, a atenção para as reações dos usuários ao se tornarem conscientes da atuação algorítmica.

O estudo parte da constatação de que não é incomum percebermos que, ao seguirmos ou deixarmos de seguir determinados contatos ou páginas, conseguimos alterar significativamente, por exemplo, nosso *feed* de notícias. Dessa forma, parece que certas ações do usuário teriam o poder de alterar o teor daquilo que é exibido no *feed*. Tendo isso em vista, ao se trabalhar com uma amostra de 3441 pessoas, o estudo de Min identificou quatro tipos de usuários de SRS:

- 1) os desengajados algorítmicamente (36,9% da amostra) – que apresentam pouca consciência dos aspectos algorítmicos envolvidos na coordenação das mídias sociais;
- 2) os curadores negativos (29,4% da amostra) – que ativamente deixam de seguir determinados contatos, procurando com isso consumir menos um tipo específico de notícia;
- 3) os curadores positivos (20,5% da amostra) – que seguem e adicionam outros contatos além dos seus, procurando consumir mais um determinado tipo de notícia;
- 4) os ativistas (13,2% da amostra) - que procuram desafiar as ações dos algoritmos curadores, escolhendo notícias que normalmente não consomem, ou mesmo alterando de propósito as configurações do seu perfil.

Como vemos, os usuários mais conscientes das intervenções algorítmicas acabam por desenvolver as mais variadas estratégias no sentido de controlá-las. Por seu turno, os menos conscientes, os “desengajados”, supostamente estariam mais expostos e sujeitos à influência da curadoria algorítmica.²³ O estudo sugere, portanto, que a maioria dos usuários dos SRS mantém certo controle sobre as informações que recebem, contrapondo-se de alguma forma à atuação curadora.

Em suma, os estudos de Bakshy *et al.* (2015) e Min (2019) sugerem que o poder de indução dos algoritmos curadores nas mídias sociais, embora não seja desprezível, não é tão significativo quanto se imaginava. Contudo, não podemos nos esquecer que tanto os motores de busca na internet quanto as mídias sociais ainda continuam categorizando e dividindo os seus usuários, o que acarreta diversas consequências, como observam Myers e Twenge (2022):

Os inúmeros grupos virtuais da internet permitem que pacifistas e neonazistas, fãs de esportes e fãs de Star Trek, teóricos da conspiração e sobreviventes de câncer se isolem com pessoas de ideias semelhantes e encontrem apoio para suas preocupações, interesses e suspeitas compartilhadas.
 [...] **Comunicar-se com pessoas que compartilham seus interesses tem muitos benefícios, mas quando as pessoas se isolam de outras opiniões, a internet pode se tornar uma câmara de eco.** (MYERS; TWENGE, 2022, p.213, tradução nossa, negrito nosso)

O conceito figurado pelo termo “**homofilia**” pode nos ajudar a entender essa situação. Este termo foi estabelecido pelos sociólogos, Paul F. Lazarsfeld e Robert K. Merton em 1954, no artigo intitulado “Friendship as a Social Process: a substantive and methodological analysis”, e refere-se à tendência das pessoas se aproximarem ou se associarem a outras que possuem características semelhantes. Dessa forma, pode-se dizer que os SRS, ao categorizarem seus usuários, favorecendo com isso o encontro entre indivíduos que pensam de modo semelhante, acabariam por facilitar a homofilia. Assim, se uma das condições necessárias para que um grupo se polarize é que seus membros pensem de modo semelhante, a homofilia, proporcionada pela

²³ Vale ressaltar que, nesse estudo, os “desengajados” eram mais velhos (mais da metade estava na faixa dos 50 anos ou mais) e eram também um pouco menos instruídos que os demais pesquisados.

categorização dos SRS, facilitaria definitivamente tal situação. Exemplos da importância da homofilia na polarização das mídias sociais são o *WhatsApp* e o *Telegram*. Vale observar que esses SRS não possuem nenhum algoritmo curador, e, mesmo assim, muitos dos seus grupos se apresentam como autênticas bolhas epistêmicas.

Adicionalmente, além de assegurar o encontro de semelhantes, e por consequência o compartilhamento de crenças e medos, as mídias sociais mantêm seus usuários em ambientes virtualmente delimitados e sob estímulo constante. O que torna a situação ainda mais dramática principalmente quando grupos virtuais disseminam notícias falsas e discursos de ódio.

Assim, desacordos podem se tornar demonizações e suspeitas podem escalar para a paranoia. [...] Com *retweets*, *feeds* de notícias personalizados e auto seleções do *buffet* de notícias, mentes semelhantes também podem alimentar umas às outras com desinformação tóxica: inverdades que, após muitas repetições, são aceitas como fato. [...] Isso é exatamente o que aconteceu durante a pandemia da COVID-19, quando desinformações e teorias da conspiração se espalharam como fogo em mato seco pela internet. (Myers; Twenge, 2022, p.213, tradução nossa)

Portanto, quando inverdades compartilhadas se transformam em fatos indiscutíveis, grupos se radicalizam, e a racionalidade das interações discursivas (asserções, argumentos e explicações) nessas redes virtuais se compromete. Por esse motivo, dada a sua proliferação, influência e falta de controle, essas bolhas epistêmicas virtuais, criadas e sustentadas por meio de crenças falsas, acabam se tornando um grave obstáculo ao nosso processo civilizatório, como vimos.

Iniciamos nosso estudo com a ideia de que o discurso racional tende aos extremos quando ocorre uma polarização grupal. Entretanto, as formulações apresentadas até este momento, bem como a análise dos casos reais de bolhas epistêmicas, revelaram um processo bem mais complexo. Nele, elementos factuais e linguísticos se influenciam mutuamente, extremando o discurso grupal e, por

consequência, conduzindo o grupo à polarização. No sentido de melhor compreender essa dinâmica, podemos sistematizá-la por meio da seguinte sequência de eventos:

1. Deve haver um estado grupal inicial em que temos um conjunto de indivíduos que pensam de modo semelhante e que se reconhecem enquanto grupo;
2. Em seguida, dada a ocorrência de um conjunto real ou imaginário de eventos catalisadores (um desafio, uma ameaça, um temor etc.), os membros desse grupo, ao se sentirem desafiados ou ameaçados, estabelecem um alvo ou um objetivo a ser alcançado (um grupo a combater, um problema a ser resolvido etc.);
3. Deve existir um sítio físico ou virtual onde os membros possam, partindo da homogeneidade de pensamento pré-estabelecida, deliberar, organizar e direcionar a razão para alcançar os objetivos propostos;
4. Finalmente, a razão se precipita na forma de um discurso mais extremado, levando o grupo a atingir o estado de polarização ou mesmo de radicalização.

4 OS MODELOS EXPLICATIVOS

Na obra citada acima, “La Démonomanie des Sorciers”, encontramos o seguinte relato de Jean Bodin:

[Li] também o relato do julgamento das bruxas de Potez, que foi comunicado a mim pelo Mestre Adrian de Fer, Tenente General de Laon, que relatou a confissão delas, de como foram transportadas [voaram] para perto de Longny, no moinho de Frenquis, ao pronunciar certas palavras, que eu não registrarei aqui, com uma vassoura ou limpador de chaminé. Elas encontraram lá as outras, cada uma com a sua vassoura, e com elas seis demônios que lá foram nomeados. E após renunciarem a Deus, beijaram os demônios em forma humana, embora extremamente horrendos de se olhar, e os adoraram. Então elas dançaram com os seus limpadores de chaminé em mãos, e finalmente os demônios copularam com as mulheres. Depois, elas pediram por pó para matar gado, tendo sido decretado que deveriam retornar lá oito dias depois, que seria segunda-feira, após o pôr do sol, e elas permaneceram lá por cerca de três horas e depois foram levadas [voaram] de volta. (BODIN, 2001, p.119)

As palavras de Bodin eram destinadas a um público bem específico: a elite letrada renascentista. Como mencionamos no capítulo anterior, a elite europeia encontrava-se polarizada em relação ao tema da feitiçaria naquela época. Esse fato levava muitos de seus membros a compartilharem e a defenderem um conjunto de crenças radicais que se concretizava na denominada doutrina da bruxaria. Por meio desse sistema de crenças e teses, seus membros conseguiam compreender e explicar eventos fantásticos, como aqueles relatados acima por Bodin, por exemplo. Tais explicações falaciosas e argumentos eram bastante persuasivos e principalmente aceitos como verdadeiros por muitos daqueles responsáveis pelo ordenamento e repressão social (nobres, juízes, juristas, membros da igreja etc.), justificando a intensa perseguição.

Por sua vez, na obra "Os Estabelecidos e os *Outsiders*", Elias e Scotson detalham um aspecto revelador do poder da fofoca depreciativa disseminada em Winston Parva pelos “membros da aldeia” (os estabelecidos), tendo como alvo os “habitantes do loteamento” (os *outsiders*):

Os mexericos discriminatórios da “aldeia”, todas as expressões de censura e desdém segredadas ou abertamente proferidas contra as pessoas do loteamento, exerciam um poder sobre elas, por mais dignas e ordeiras que fossem em sua conduta, porque parte delas mesmas, sua própria consciência, concordava com a má opinião que os “aldeões” tinham de seu bairro. Era essa concordância silenciosa que lhes paralisava a capacidade de retaliar e de se afirmar. Elas se sentiam envergonhadas quando alguém se referia ao grupo a que pertenciam por um termo depreciativo, ou quando eram acusadas, direta ou indiretamente, de malfeitos e falhas que, na verdade, só eram encontrados em seu grupo na “minoría dos piores”. (ELIAS; SCOTSON, 2000, p.131)

Por que os *outsiders* aceitavam os comentários maldosos e falaciosos como se fossem verdadeiros para todo o grupo, mesmo sabendo que muitas vezes se referiam a “minoría dos piores”? Por que essa fofoca depreciativa, essa hipótese explicativa falsa, conforme mostramos no último capítulo, tinha tanto poder?

Finalmente, durante a pandemia da COVID-19 no Brasil, inúmeros grupos virtuais se radicalizaram em torno de crenças negacionistas e anticientíficas. Em muitos desses grupos, propagava-se a ideia de que a doença era apenas uma gripe comum, sugerindo que medidas de prevenção, como o uso de máscaras, o distanciamento social ou mesmo a vacinação, eram desnecessárias. Propagava-se também o uso de tratamentos não comprovados cientificamente, como a hidroxicloroquina e a ivermectina, de efeitos colaterais graves para a saúde. Ademais, nas mídias sociais, muitas explicações fantasiosas e falaciosas que sustentavam esse estado de coisas tinham como origem as próprias autoridades governamentais. Por que explicações falaciosas, como as negacionistas, têm o poder de influenciar o comportamento de parte da população, suscitando, como presenciamos, consequências tão trágicas e fatais?

Em suma, como podemos observar nos exemplos de bolhas epistêmicas apresentados acima, por que explicações obviamente falaciosas podem ser tão persuasivas? Ou ainda, nos termos do nosso trabalho, por que a “razão comprometida” pode ser tão efetiva?

No sentido de se compreender o poder da explicação falaciosa, há a necessidade de se definir alguns conceitos com mais precisão. Para tanto, iniciaremos com a apresentação do conceito figurado pelo termo “**explicação**”. Esta conceituação vem sendo discutida principalmente no campo da filosofia da ciência desde meados do século XX. Devido à sua importância e influência, dois modelos se destacaram na

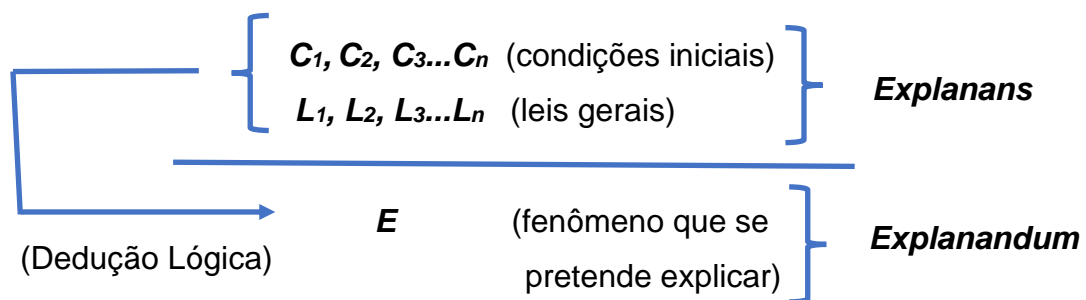
tentativa de conceituação: o **modelo nomológico-dedutivo** e o **modelo de relevância estatística**. Embora esses modelos tenham sido desenvolvidos visando à explicação de fenômenos nas ciências exatas e biológicas, nada impede que sejam aplicados no campo das ciências humanas ou mesmo na compreensão de eventos do nosso dia a dia.

4.1 O modelo nomológico-dedutivo

Este constructo, também conhecido como modelo de cobertura por leis, foi um dos primeiros modelos a colocar no centro das discussões filosóficas a análise e o estudo das estruturas envolvidas em um processo explicativo. Ele foi desenvolvido a partir de várias obras por Carl Hempel e Paul Oppenheim. A ideia é simples: uma explicação poderia ser modelada por meio de um argumento dedutivo:

A explicação aqui descrita pode ser considerada um argumento segundo o qual um fenômeno que se pretende explicar, o fenômeno do *explanandum*, ocorre em virtude de certos fatos explicativos [*explanans*]. Estes se classificam em dois grupos: (i) fatos particulares [condições iniciais] e (ii) uniformidades expressas por meio de leis gerais (HEMPEL, 1965, p. 336, tradução nossa).

Assim, de modo esquemático, uma explicação poderia ser representada por um argumento dedutivo da seguinte forma:



Para compreendermos o funcionamento do modelo na prática, analisemos o seguinte exemplo ilustrativo que foi inspirado em um exercício proposto por Halliday *et al.* (2022, p.126). Tarzan, ao saltar de um penhasco pendurado na extremidade de

um cipó, como habitualmente fazia, foi surpreendido com o seu rompimento justamente quando se encontrava no ponto mais baixo da trajetória. Nessa situação, gostaríamos de saber: por que o cipó se rompeu?

Ao modelar a situação, devemos primeiramente definir os referenciais teóricos e conceituais que usaremos. No nosso caso, escolhemos a mecânica newtoniana como o conjunto de teorias e conceitos que delimitarão o argumento que figura o modelo. Dessa forma, as premissas e a conclusão desse argumento representarão informações, conceitos e leis gerais alinhados a tal enquadramento teórico. Nesses termos, dada a verdade das premissas, a conclusão, ou melhor, o *explanandum*, deve ser igualmente verdadeiro. Portanto, caso o argumento assim construído seja legítimo,²⁴ podemos dizer que o evento foi explicado.

No exemplo, na construção do argumento que procurará explicar o rompimento do cipó, a conclusão (*explanandum*) que queremos provar é:

A tensão gerada pelo movimento do herói, no ponto mais baixo da trajetória, é maior que a tensão máxima que o cipó pode suportar, ou seja: $T_{\text{herói}} > T_{\text{máxima-cipó}}$

Como primeira premissa do argumento, devemos elencar todas as condições iniciais conhecidas. Assim, suponhamos que conhecemos as seguintes informações:

- l = comprimento do cipó = 18 m
- h = altura entre o alto do penhasco e o ponto mais baixo da trajetória = 3,2 m
- m = massa do herói = 80 kg
- g = aceleração da gravidade = 9,8 m/s²
- $T_{\text{máxima-cipó}}$ = tensão máxima suportada pelo cipó = 950 N

A segunda premissa do argumento envolverá duas leis gerais, a saber:

- i. Segunda Lei de Newton. No ponto mais baixo do movimento, a resultante das forças (tensão no cipó devido ao movimento menos o peso do herói) é igual à massa vezes a aceleração centrípeta do herói, o que nos conduz à seguinte

²⁴ Um argumento é considerado legítimo ou sólido quando é formalmente válido (ou seja, se suas premissas fossem verdadeiras, a conclusão também seria necessariamente verdadeira), e suas premissas são de fato verdadeiras.

fórmula: (1) $T_{\text{herói}} - m \cdot g = m \cdot v^2 / l$; onde v é a velocidade tangencial do herói no ponto mais baixo da trajetória;

- ii. Lei de Conservação de Energia. Ou seja, ao longo de todo o movimento, a soma da energia potencial com a energia cinética deve manter-se constante, o que nos conduz, após algum manuseio algébrico, à seguinte fórmula: (2) $m \cdot g \cdot h = m \cdot v^2 / 2$; onde v é a velocidade tangencial do herói no ponto mais baixo da trajetória.

Do manuseio algébrico de (1) e (2), chegamos à seguinte fórmula:

$$(3) T_{\text{herói}} = m \cdot g (1 + 2 h / l)$$

Finalmente, substituindo os valores iniciais na fórmula (3), chegamos ao seguinte valor: $T_{\text{herói}} = 1062,75 \text{ N}$. Assim, ao confirmarmos a verdade da conclusão do argumento, uma vez que $T_{\text{herói}} (1062,75 \text{ N})$ é maior que a $T_{\text{máxima-cipó}} (950 \text{ N})$, explicamos o evento em questão. Ou seja, no ponto mais baixo da trajetória, o cipó se rompeu porque foi submetido a uma tensão maior do que a máxima suportada.

No modelo nomológico-dedutivo, as leis gerais devem possuir um caráter universal. Contudo, muitas vezes, essas normas são de natureza estatística. Tendo isso em vista, Hempel propõe um segundo modelo: o estatístico-indutivo. Nesta reformulação, as leis gerais passam a ter um caráter probabilístico, o que confere ao constructo argumentativo uma natureza indutiva. Tal natureza implica, ao final do processo argumentativo, um *explanandum* de caráter probabilístico. Não discutiremos o constructo estatístico-indutivo, uma vez que o modelo nomológico-dedutivo serve bem ao nosso intuito atual que é a análise geral das ideias desenvolvidas por Hempel e suas consequências.

Vale notar que uma característica que confere robustez ao modelo nomológico-dedutivo é a previsibilidade de eventos que ele proporciona. Isto é, uma vez conhecidos os valores das condições iniciais, sempre que precisarmos explicar o porquê de um determinado cipó ter se rompido em circunstâncias semelhantes às do exemplo, basta calcular a tensão no ponto mais baixo do movimento, como fizemos.

Entretanto, ainda que robusto, o modelo nomológico-dedutivo apresenta algumas deficiências. Segundo Salmon (2013), as principais críticas podem ser apresentadas na forma de contraexemplos. Assim, de acordo com esse modelo, caso não conhecêssemos o peso do nosso herói, poderíamos calculá-lo a partir das mesmas condições iniciais²⁵ e leis gerais²⁶. Todavia, não parece intuitivo pensar que uma correta explicação para o peso atual do nosso personagem esteja relacionada ao comprimento do cipó ou à tensão no ponto mais baixo da trajetória, por exemplo. No modelo nomológico-dedutivo, os elementos devem ser alocados no constructo conforme uma certa direção. Dessa forma, ao reunir os elementos naquela disposição específica, a modelagem do exemplo explica apenas o rompimento do cipó. Ou seja, a massa do personagem apresenta uma função e uma posição rígida naquela modelagem. Ao se estudar casos semelhantes a esse, confirma-se um fato há muito conhecido pelos estudiosos do tema: se, em alguns casos, um argumento dedutivo permite a “simetria dedutiva”, ou melhor, se é possível provar a verdade de uma das premissas a partir da verdade da conclusão e das demais premissas; então, o modelo nomológico-dedutivo, enquanto um argumento dedutivo, permitiria igualmente a construção de “simetrias explicativas”. Contudo, como mostra a explicação do peso do herói a partir do comprimento do cipó, tal simetria explicativa conduz a uma construção modelar incorreta. Vejamos agora este contraexemplo:

O homem e a pílula. Um homem, ao explicar o seu fracasso em engravidar ao longo do ano passado, alega que consumiu regularmente as pílulas anticoncepcionais da sua esposa e que qualquer homem que toma regularmente anticoncepcionais orais nunca engravidará. (SALMON, 2013, p.350, tradução nossa)

Segundo o modelo nomológico-dedutivo, não há nada de errado com essa explicação. Entretanto, é notório que os elementos do *explanans* (isto é, o fato do homem consumir regularmente pílulas anticoncepcionais e de que todo homem que as consumir nunca engravidará) são irrelevantes para a explicação do fracasso do homem em engravidar. Por fim, examinemos este outro contraexemplo:

²⁵ É claro que, nesse caso, a tensão ($T_{\text{herói}}$) no ponto mais baixo da trajetória passa a ser uma das condições iniciais.

²⁶ Poderíamos usar a seguinte fórmula no cálculo da massa do herói: $m = T_{\text{herói}} / (g (1 + 2 h / l))$.

A mancha de tinta. No tapete, próxima à mesa do escritório do professor Jones, há uma grande mancha preta. Como ele explicaria isso? Ontem, próximo ao canto da sua mesa, encontrava-se um frasco de tinta preta aberto. Acidentalmente, ele bateu com o cotovelo no frasco, que, ao cair no chão, derramou tinta sobre o tapete. (SALMON, 2013, p.350, tradução nossa)

No contraexemplo, a explicação para a existência da mancha no tapete não precisou seguir o padrão exigido pelo modelo de Hempel. Ou seja, não foi necessário evocar nenhuma lei geral para se compreender o evento.

Dados os problemas apresentados acima pelos contraexemplos, Salmon (2006) propõe uma nova formulação que se dispõe a superar as falhas do modelo de Hempel. Esse novo modelo visa impedir tanto as assimetrias explicativas indevidas quanto a consideração de possíveis fatos irrelevantes na explicação do evento.

4.2 O modelo de relevância estatística

Proposto por Salmon (2006), este modelo não exibe a forma de um argumento. A modelagem apresenta-se enquanto um conjunto de sentenças organizadas em uma configuração específica:

Uma explicação RE [segundo um modelo de Relevância Estatística] consiste numa probabilidade *a priori*, uma partição homogênea relevante em relação ao atributo em questão, as probabilidades *a posteriori* do atributo nas células da partição e uma declaração da localização do indivíduo em uma determinada célula da partição:

- $P(B/A) = p$;
- $P(B/A.C1) = p1$; $P(B/A.C2) = p2$; . . .
- A.C1, A.C2, . . . constitui uma partição homogênea relevante de A;
- b é um membro de A.Ck.

(SALMON, 2006, p.65, tradução nossa)

Para efeito de comparação, vamos formular a explicação do exemplo do rompimento do cipó, de acordo com o modelo de relevância estatística. Para começar, continuamos a considerar o contexto da mecânica newtoniana como nosso campo teórico de trabalho e análise. Diante disso, de acordo com o modelo de Salmon, podemos afirmar que o cipó que se rompeu (c_{romp}) pertence à classe dos cipós da floresta (c_{floresta}). Em seguida, dividimos essa classe, c_{floresta} , de acordo com fatores

que consideramos relevantes para a ocorrência do evento. Para tanto, sugerimos como fator relevante para o rompimento do cipó, a submissão deste a uma tensão maior do que a máxima suportada ($T_{\text{maior que a máxima}}$).²⁷

Tendo isso em mente, definimos as partições, $C_{\text{floresta}}(T_{\text{maior que a máxima}})$ e $C_{\text{floresta}}(\sim(T_{\text{maior que a máxima}}))$. Elas representariam respectivamente a subclasse dos cipós que estariam submetidos a uma tensão maior do que a máxima e a subclasse em que estes estariam submetidos a uma tensão menor ou igual à máxima permitida. Finalmente, calculamos as probabilidades *a priori* e *a posteriori* envolvidas no evento por meio de uma análise estatística:

- (1) $P(\text{cipó se romper} / C_{\text{floresta}}) = 0,07$. Esta é a probabilidade (*a priori*) de qualquer cipó da floresta se romper quando submetido a alguma tensão. Pode-se calculá-la de várias maneiras. Um dos métodos envolve a realização de observações ou cálculos contínuos ao longo do tempo. Por exemplo, em um caso hipotético, anotamos o número de vezes (**m**) em que Tarzan, Jane e Chita utilizaram, nas mais variadas situações, os cipós da floresta para se locomoverem em um período de um ano. Igualmente anotamos quantas vezes (**n**) algum cipó se rompeu naquelas ocasiões. O simples cálculo da razão **n/m** forneceria, nesse caso, um valor adequado para a probabilidade *a priori*. Vale notar que o valor baixo dessa probabilidade explica a razão pela qual o herói e sua família usam habitualmente esse “meio de transporte”.

- (2) $P(\text{cipó se romper} / C_{\text{floresta}}(T_{\text{maior que a máxima}})) = 0,90$. Este valor indica a probabilidade (*a posteriori*) de qualquer cipó da floresta se romper, desde que seja submetido a uma tensão maior do que a máxima permitida. Tanto essa probabilidade quanto a próxima foram obtidas a partir de um ensaio realizado durante a modelagem do evento.²⁸

²⁷ No caso, no sentido de se simplificar o entendimento do modelo, escolhemos apenas uma condição relevante para a ocorrência do evento.

²⁸ Por exemplo, nesse ensaio hipotético, em média, 9 de cada 10 cipós submetidos a uma tensão maior do que a máxima se romperam. Por outro lado, em média, 1 de cada 200 cipós submetidos a uma tensão menor ou igual à máxima se rompeu.

(3) $P(\text{cipó se romper} / c_{\text{floresta}} \cdot (\sim(T_{\text{maior que a máxima}}))) = 0,005^{29}$. Indica a probabilidade (*a posteriori*) de um cipó qualquer da floresta se romper, desde que submetido a uma tensão menor ou igual à máxima permitida.

A modelagem do exemplo apresenta-se como um conjunto composto por quatro sentenças. Ou seja, (1), (2) e (3) acima, mais a seguinte sentença:

(4) (c_{romp}) é um membro da partição $c_{\text{floresta}} \cdot (T_{\text{maior que a máxima}})$.

Esta última sentença representa o fato de que o cipó usado pelo herói pertence à subclasse daqueles cipós da floresta que são submetidos a uma tensão maior que a máxima suportada. Uma vez que a probabilidade $P(\text{cipó se romper} / c_{\text{floresta}} \cdot (T_{\text{maior que a máxima}}))$ (0,90) é bem maior que a probabilidade *a priori* $P(\text{cipó se romper} / c_{\text{floresta}})$ (0,07) e que a probabilidade $P(\text{cipó se romper} / c_{\text{floresta}} \cdot (\sim(T_{\text{maior que a máxima}})))$ (0,005), podemos afirmar que a tensão $T_{\text{maior que a máxima}}$ é um fator relevante para o evento (o rompimento do cipó). Desse modo, ao ser submetido a uma tensão maior que a máxima permitida, seria alta a probabilidade de um cipó qualquer daquela floresta se romper. Portanto, a partir da análise em conjunto das quatro sentenças, conseguimos compreender o porquê do rompimento do cipó, ou melhor, conseguimos explicar o evento.

O modelo de Salmon também se mostra eficaz quando confrontado com os contraexemplos mencionados. Por exemplo, utilizando-se os mesmos dados, não seria possível obter probabilidades *a posteriori* adequadas que explicassem, por exemplo, o peso do nosso herói, dado que as informações disponíveis não se apresentam como fatores relevantes para a determinação de sua massa³⁰. Do mesmo modo, no caso do exemplo do **homem e a pílula**, o fato de ter regularmente tomado pílulas anticoncepcionais não é um fator relevante para a não gravidez do homem. Por fim, como o modelo não exige leis gerais para a sua construção, poderíamos encontrar um conjunto adequado de sentenças que explicasse **a mancha de tinta** no

²⁹ Mesmo que submetido a uma tensão menor ou igual à máxima permitida, o cipó ainda poderia se romper devido a um problema estrutural qualquer. Por exemplo, ele poderia estar muito seco, ou até ter sido sabotado por algum inimigo.

³⁰ Nesse caso, fatores relevantes poderiam ser, por exemplo, a estrutura física do personagem ou a sua dieta alimentar.

tapete do professor Jones. Entretanto, o próprio Salmon (2013) alerta-nos que, embora o modelo de relevância estatística apresente melhorias em relação à ideia geral do modelo de Hempel, há um problema em sua construção. Ou seja, o modelo de relevância estatística baseia-se em cálculos probabilísticos, o que lhe confere um caráter correlacional e não causal. Por esse motivo, devemos ter sempre em mente que, ao usarmos esse modelo, estamos trabalhando com relevâncias de natureza estatística em vez de relevâncias causais.

4.3 Os aspectos pragmáticos envolvidos numa explicação

Além dos problemas apresentados acima pelo modelo nomológico-dedutivo e pelo modelo de relevância estatística, Faye (2014) propõe um conjunto de críticas que transcende tanto as dificuldades observadas na aplicação do primeiro modelo a determinados contraexemplos, quanto aos problemas que emergem da natureza probabilística do processo explicativo quando se emprega o segundo constructo.

As generalizações empíricas geralmente contêm cláusulas *ceteris paribus* que as tornam verdadeiras apenas quando certas idealizações são cumpridas; entretanto, tais condições idealizadas nunca são alcançadas neste mundo. Além disso, nenhuma declaração nômica fundamental se aplica diretamente ao mundo; quando as declarações nômicas são verdadeiras, elas são verdadeiras para um modelo idealizado. Enunciados nômicos, como as leis de Newton, nunca se referem a nenhum objeto em particular; eles apenas expressam como várias propriedades estão inter-relacionadas. Os modelos representam sistemas concretos enquanto podem ser retratados de acordo com certos padrões interpretativos. (FAYE, 2014, p.186, tradução nossa)

Uma vez que as leis gerais são verdadeiras apenas sob determinadas condições ideais, aplicá-las na elucidação de eventos reais exige do explicador certa discricionariedade na interpretação dos elementos que comporiam uma explicação. Não se pode, portanto, negligenciar a influência das circunstâncias em que se encontra um explicador nem desconsiderar suas intenções e interesses.

[Os] físicos selecionam uma lei de cobertura específica com base em critérios pragmáticos. Nenhuma descrição teórica se aplica diretamente (ou seja, sem interpretação) ao mundo. Portanto, esses

critérios pragmáticos não podem ser considerados irrelevantes, mas são essenciais para se compreender a noção de explicação científica. [...] Um tratamento pragmático e naturalista da explicação concentre-se na compreensão humana. [...] A abordagem naturalista deve incluir todos os tipos de explicações, especialmente, as que dependem de informações sobre motivos ou significados. [...] Os naturalistas abordam essa prática como expressão da compreensão do explicador não apenas do conteúdo do que é requerido na explicação, mas também de sua própria compreensão dos fatos que fornecem a informação desejada. (FAYE, 2014, p.186–187, tradução nossa)

Em suma, nenhum estudo científico seria significativo se suas explicações, além dos óbvios aspectos semânticos, não levassem em consideração os denominados critérios pragmáticos, ou seja: as circunstâncias que envolvem os enunciados e a interpretação dos elementos que os compõem. Dentre esses elementos, destacam-se os fatores relevantes:

Com certeza, a semântica tem algum papel na determinação de uma resposta relevante, entretanto ela não consegue realizar todo o trabalho sozinha, pois o contexto também é necessário para determiná-la. Em outras palavras, sem especificar um contexto, não se pode saber que tipo de resposta seria relevante (FAYE, 2014, p.188–189, tradução nossa).

Diante das considerações apresentadas acima, como poderíamos analisar os aspectos contextuais e intencionais dos agentes envolvidos em uma interação explicativa?

[O] termo "explicação" às vezes refere-se a entidades linguísticas — isto é, coleções de declarações sobre fatos — e outras vezes a entidades não linguísticas — ou seja, esses mesmos fatos. Quando pensamos em termos da atividade humana de explicar algo para alguém ou grupo de pessoas, estamos considerando o comportamento linguístico. Explicar algo a alguém envolve proferir ou escrever declarações. (SALMON, 2013, p.360, tradução nossa)

Entender uma explicação como uma atividade humana que se configura por meio de um determinado comportamento linguístico abre novas e interessantes perspectivas de estudo. Será que, a partir da análise da linguagem ou, mais precisamente, da análise das estruturas linguísticas de um discurso explicativo,

conseguiríamos avaliar a influência do contexto e das intenções dos envolvidos em sua expressão? A resposta é sim. Para tanto, já dispomos há algum tempo de um instrumento bastante adequado: a pragmática linguística. Ramo linguístico de desenvolvimento recente, a pragmática linguística apresenta-se como uma classe de teorias e estudos que em conjunto com a sintática e a semântica formam a clássica tríade da análise discursiva. *Grosso modo*, a sintática é o estudo da relação entre os diversos elementos linguísticos de uma determinada proposição. Por seu turno, a semântica relaciona esses elementos com o que os denota (suas referências). Por fim, a pragmática estuda a utilização desses elementos por usuários e intérpretes nas mais diversas situações linguísticas. Vale notar que, além das circunstâncias evidentes de fala, a pragmática linguística também se ocupa dos desejos, intenções e crenças dos indivíduos envolvidos em uma dada situação dialógica.

No estudo de modelos explicativos, as análises de natureza linguístico-pragmática ganharam atenção de estudiosos e pesquisadores a partir do final do século XX, inspirando o desenvolvimento de novos modelos que incorporaram explicitamente parâmetros relacionados à relevância e à escolha de teorias em suas arquiteturas. Neste contexto, o modelo proposto por van Fraassen (2007) merece destaque. Devido às suas peculiaridades, que se alinham com os objetivos do nosso estudo, o modelo de van Fraassen será detalhado e analisado mais adiante. Entretanto, antes de prosseguirmos, é necessário esclarecer alguns pontos.

Na apresentação dos casos ilustrativos de bolhas epistêmicas, observamos como os grupos compostos por indivíduos que pensam de modo semelhante podem se polarizar e até se radicalizar ao deliberarem. Esse fato sugere que, conforme as crenças (e, por conseguinte, os argumentos e as explicações) se tornam mais radicais, algum tipo de mutação contextual também deve estar ocorrendo no interior desses grupos. Nessa situação, no que se refere especificamente ao discurso explicativo, um determinado evento, até então interpretado sob uma perspectiva particular, poderia, em um contexto de polarização, ser agora explicado segundo diferentes teorias e aspectos relevantes. Ademais, a eficácia dos discursos explicativos falaciosos dentro dessas bolhas epistêmicas é notória. O que indica que

a escolha deliberada de perspectivas de mundo, nesses grupos, muitas vezes ocorre sem maiores considerações no nível semântico. Diante desses aspectos, não seriam os modelos explicativos com ênfase pragmática, ou seja, aqueles em que são determinantes os contextos que cercam as interações linguísticas, a despeito das considerações de cunho semântico, os mais apropriados aos nossos objetivos? Tendo em vista essa questão, assim como outros aspectos igualmente importantes, utilizaremos como ponto de partida analítico a teoria da explicação pragmática de Bas C. van Fraassen. Tal formulação possibilita a construção de um modelo explicativo que, entre outras qualidades, incorpora elementos contextuais e aspectos relevantes na forma de parâmetros em sua arquitetura. Tais características fazem deste modelo um instrumento teórico bastante adequado aos nossos propósitos, como veremos.

5 A TEORIA PRAGMÁTICA DA EXPLICAÇÃO DE van FRAASSEN

A morte do piloto Ayrton Senna em 1994 em Ímola causou enorme comoção. Várias explicações sobre o que teria causado o óbito do piloto foram cogitadas pelos meios de comunicação naquela época. *Grosso modo*, elas se agrupavam em torno das seguintes causas:

- o traçado inseguro da curva Tamburello;
- uma emenda na barra de direção que teria se rompido instantes antes do acidente;
- a perfuração craniana fatal devido a um pedaço da barra de suspensão que se soltou durante o acidente;
- a intensa pressão psicológica por melhores resultados.³¹

Durante os dias que se seguiram, vimos diversos especialistas apresentarem, dentro de seus ramos de estudo, inúmeras explicações que se sustentavam em uma ou outra das causas elencadas acima. Nessas ocasiões, era como se o mecânico dissesse, "se a barra de direção não tivesse se rompido, não haveria acidente e nesse caso...". Ou no caso do médico, "se não ocorresse a hemorragia devido à perfuração craniana, Ayrton poderia ter sobrevivido". Ou ainda no caso do psicólogo, "se não ocorresse a pressão psicológica por resultados, não haveria nenhum motivo para um comportamento mais arriscado, o que...". Essa série de contrafactuais demonstra uma característica essencial, projetistas de circuitos, mecânicos especializados, médicos, psicólogos e ex-pilotos estavam todos executando um mesmo movimento: **ao explicar, apresentavam, segundo seus pontos de vista, aspectos objetivamente relevantes para o entendimento do evento**. Na obra "A Imagem Científica", van Fraassen (2007) ressalta a importância desses aspectos, bem como defende a ideia de que uma explicação sempre visa à resolução de um tipo de questão, a saber, uma questão-por-quê (*why-question*) – "Por que (é o caso que) p?".

³¹ O fraco desempenho no campeonato teria levado tanto o piloto quanto a sua equipe a adotar um comportamento mais arriscado do que o normal, o que teria motivado, na busca por uma dirigibilidade melhor do veículo, a emenda na barra de direção.

Uma explicação não é o mesmo que uma proposição, ou um argumento, ou uma lista de proposições; ela é uma resposta. [...] Uma explicação é uma resposta para uma questão-por-quê. Assim, uma teoria da explicação deve ser uma teoria das questões-por-quê. (VAN FRAASSEN, 2007, p.237)

Hempel (1965) já afirmara tal fato:

Uma explicação científica pode ser considerada uma resposta a uma questão-por-quê, como: "Por que os planetas se movem em órbitas elípticas tendo o Sol em um dos focos?", "Por que a lua parece muito maior quando está próxima do horizonte do que quando está no alto do céu?", "Por que o aparelho de televisão do Ranger VI falhou?", "Por que os filhos de pais com olhos azuis terão sempre olhos azuis?", "Por que Hitler entrou em guerra contra a Rússia?". (HEMPEL, 1965, p.334, tradução nossa)

Mais precisamente, na situação do exemplo, é importante observar como todas as explicações dos especialistas visam responder à mesma questão-por-quê — **Por que Ayrton Senna faleceu naquele dia?** Segundo van Fraassen, na enunciação de uma questão-por-quê, a sua possível interpretação encontra-se, no momento de sua expressão, delimitada pelas circunstâncias e pelas informações disponíveis bem como pelos interesses dos envolvidos. À medida que esses aspectos mudam, a resposta à questão-por-quê – ou melhor, a explicação – também muda.

[Uma] explicação frequentemente consiste em listar fatores salientes, que indicam um relato completo de como o evento aconteceu. O efeito disso é o de eliminar diversas hipóteses alternativas sobre como esse evento realmente se deu e/ou eliminar algum embaraço que diz respeito a como o evento poderia se ter dado. Mas a saliência é dependente de contextos, e a escolha dos fatores corretos e “importantes” depende do âmbito de alternativas consideradas naquele contexto. (VAN FRAASSEN, 2007, p.229)

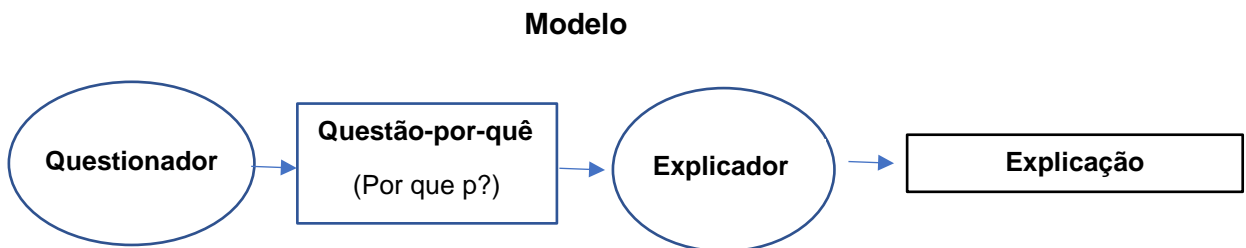
Vale ressaltar, portanto, que uma explicação se constitui por meio da escolha de um subconjunto adequado de fatores explicativos sobre um determinado universo de possibilidades, o que implica a eliminação de outros tantos subconjuntos possíveis como veremos em maior detalhe adiante.

[O] advogado quer esse tipo de eliminação de hipóteses sobre a morte que sejam relevantes para a consideração jurídica; o construtor de carrocerias, uma eliminação de hipóteses sobre os defeitos estruturais

ou as limitações estruturais em diversos tipos de esforços mecânicos. Em outras palavras, o contexto determina a relevância de uma forma que vai bem além da relevância estatística sobre a qual nossas teorias científicas dão informação.

Isso poderia não ser importante se nós não estivéssemos preocupados em descobrir exatamente como ter uma explicação vai além de meramente ter uma teoria aceitável sobre o domínio de fenômenos em questão. Mas esse é exatamente o centro de nossas preocupações. (VAN FRAASSEN, 2007, p.229)

Na linha argumentativa de van Fraassen, um ato explicativo pode ser pensado como o resultado da aplicação de um determinado modelo. Este se caracteriza como uma interação entre um questionador real ou fictício³², que formula uma questão-por-quê, e um explicador, que procura respondê-la, de acordo com o esquema a seguir.



Considerar tanto as intenções dos envolvidos quanto os contextos que permeiam a enunciação da questão-por-quê e da respectiva explicação, significa direcionar a análise no sentido do universo da pragmática linguística. De acordo com Woodward e Ross:

Podemos afirmar que uma teoria da explicação contém elementos "pragmáticos" se, conforme a teoria, (i) tais elementos necessitam de uma referência irreduzível aos fatos sobre os interesses, crenças ou outros aspectos psicológicos daqueles que estão fornecendo ou recebendo a explicação e/ou (ii) uma referência irreduzível ao "contexto" no qual a explicação está sendo realizada. (WOODWARD; ROSS, 2021, tradução nossa)

³² De acordo com o modelo, por exemplo, quando um professor explica uma determinada matéria, sua explicação não é motivada na maioria das vezes por um questionador real. Neste caso, ele articula suas explicações como se estivesse respondendo a um questionador fictício. Em alguns momentos, inclusive, ele torna explícita a figura desse questionador ao dizer retoricamente: ***“mas alguém pode perguntar: por que essa esfera chegaria com essa velocidade nesta posição? Eu explico!”***.

Ou ainda, conforme van Fraassen:

Os fatores fundamentais na situação linguística, concebidos pragmaticamente, são o falante ou usuário, a entidade sintática (sentença ou conjunto de sentenças) proferida ou apresentada, a audiência, e as circunstâncias factuais. Qualquer fator que esteja relacionado com o falante ou com a audiência é um fator pragmático; e é um fator contextual se, além disso, pertencer especificamente aquela situação linguística particular. (VAN FRAASSEN, 2007, p.165)

Assim, seja nos discursos explicativos científicos ou naqueles que utilizamos frequentemente no nosso dia a dia, os distintivos pragmáticos (crenças, intenções e desejos dos envolvidos e o contexto) influenciam as escolhas em uma pré-avaliação tanto de teorias disponíveis quanto do que é relevante naquela circunstância. Por exemplo, especificamente no caso das teorias científicas:

[S]eria um erro desconsiderar os modos nos quais a avaliação [das teorias científicas] é influenciada por fatores contextuais. Esses fatores são trazidos a situação pelo cientista, de sua condição pessoal, social e cultural. É um erro pensar que os termos nos quais uma teoria científica é avaliada sejam puramente [...] [objetivos] e que não tenham nada a ver com outro tipo de avaliação ou com as pessoas e circunstâncias envolvidas.

A aceitação de teorias possui uma dimensão pragmática. (VAN FRAASSEN, 2007, p.160)

Considerações sobre a importância do contexto e das circunstâncias de fala não pretendem colocar em segundo plano os possíveis aspectos teóricos, ou atenuar a importância das contrapartidas factuais aos discursos explicativos. A grande novidade na construção analítica empreendida por van Fraassen é a equiparação entre esses elementos:

A discussão da explicação tomou o caminho errado desde o começo, quando a explicação foi concebida como uma relação como a descrição: uma relação entre teoria e fato. Realmente, trata-se de uma relação de três termos, entre teoria, fato e contexto. Não é de se admirar que nenhuma relação simples entre teoria e fato jamais tenha conseguido se ajustar a mais que uns poucos exemplos! Ser uma explicação é algo essencialmente relativo, pois uma explicação é uma resposta. (No mesmo sentido, ser uma filha é algo relativo: toda mulher é uma filha, e toda filha é uma mulher; todavia, ser uma filha não é o mesmo que ser uma mulher). (VAN FRAASSEN, 2007, p.273, negrito nosso)

Por consequência, em um dado contexto, teorias envolvidas na explicação de um evento seriam inúteis em outras circunstâncias. Por exemplo, teorias referentes ao comportamento da fisiologia humana ou sobre o socorro e o tratamento de traumas encefálicos em acidentes automobilísticos têm um papel muito relevante nas explicações que esperamos ouvir de médicos. No entanto, quando se procura explicar as causas de uma possível falha mecânica fatal, essas teorias são inúteis.

Por fim, ao constatar a influência da pragmática na construção de um dado discurso explicativo, van Fraassen infere um resultado surpreendente:

Assim, a explicação científica não é ciência (pura), mas uma aplicação da ciência. Ela é o uso da ciência para satisfazer alguns de nossos desejos; e esses desejos são bem específicos em um contexto específico, mas eles são sempre desejos de informação descritiva. (Lembremos que toda filha é uma mulher.) O conteúdo exato do desejo e a avaliação do quanto ele é satisfeito, isso varia de contexto para contexto. Não se trata de um desejo único, o mesmo em todos os casos, por um tipo muito especial de coisa, mas, ao contrário, em cada caso, um desejo diferente por alguma coisa de um tipo bem conhecido. (VAN FRAASSEN, 2007, p.274, negrito nosso)

De acordo com Dutra (1993), filósofos como Popper, Hempel, Reichenbach e Salmon conceberam a tecnologia (a aplicação científica) como autônoma em relação à ciência pura e mantiveram a visão de que a tarefa da ciência é explicar. Tais perspectivas tornam a concepção de van Fraassen ainda mais significativa e inovadora. Ou seja, ao contrário do que esses filósofos defendiam, van Fraassen propõe que uma explicação científica é essencialmente uma questão de ciência aplicada, assim como os desenvolvimentos tecnológicos.

Mas afinal de contas, em que consiste o modelo de van Fraassen? Iniciemos pela análise do seguinte enunciado:

Por que Adão comeu a maçã?

Ao proferirmos essa questão-por-quê, conforme a nossa intenção, podemos incitar a interpretação do nosso interlocutor a três direções distintas (de acordo com os grifos abaixo):

- (1) Por que **Adão** comeu a maçã?
- (2) Por que Adão **comeu** a maçã?
- (3) Por que Adão comeu **a maçã**?

No enunciado (1), nossa atenção é direcionada para o autor da ação, e um questionamento semelhante nesse sentido poderia ser:

- (1') Por que **Adão, e não Eva**, comeu a maçã?

No enunciado (2), nossa atenção é direcionada para a ação do sujeito, e um questionamento semelhante poderia ser:

- (2') Por que Adão **comeu e não ofereceu** a Eva a maçã?

Por fim, no enunciado (3), nossa atenção é direcionada para o objeto que sofreu a ação, e um questionamento semelhante poderia ser:

- (3') Por que Adão comeu **a maçã e não o morango**?

Cada par de questões $\{(1);(1')\}$, $\{(2);(2')\}$ e $\{(3);(3')\}$ estabelece diferentes questionamentos e, por consequência, solicita diferentes respostas explicativas. Diante de tais peculiaridades, quando da enunciação de uma questão-por-quê, segundo van Fraassen, podemos observar uma determinada relação entre o que denominamos de **tema ou tópico da questão (P)** e uma determinada **classe de contraste (X)**, em que o primeiro é um elemento da segunda. Exemplificando:

- para a questão (1), temos como tema da questão (**P₁**) = Adão; e como classe de contraste (**X₁**) = {Adão, Eva, Caim, Abel};

- para a questão (2), temos como tema da questão (**P**₂) = comeu; e como classe de contraste (**X**₂) = {comeu, ofereceu, arremessou};
- para a questão (3), temos como tema da questão (**P**₃) = maçã; e como classe de contraste (**X**₃) = {maçã, morango, banana}.

Assim, podemos provisoriamente estabelecer uma primeira forma geral para uma questão-por-quê:

Por que (**P**) em contraste com (**X**)? Onde (**X**) é a classe de contraste e (**P**) é o tópico ou tema da questão.

Além do tópico ou tema da questão (**P**) e da classe de contraste (**X**), na busca de uma completa caracterização de uma questão-por-quê, mais um fator deve ser levado em consideração: a relação de relevância explicativa (**R**). Essa relação expressa-se por um conjunto apropriado de eventos que possivelmente conduziram à ocorrência do tema da questão (**P**). Assim, para o enunciado (1) poderíamos ter as seguintes relações de relevância:

- **R**₁ = eventos relacionados à disposição espacial do agente em relação ao objeto;
- **R'**₁ = eventos relacionados à estrutura hierárquica do grupo social do agente.

Por sua vez, para o enunciado (2):

- **R**₂ = eventos relacionados aos estados fisiológicos do agente;
- **R'**₂ = eventos relacionados aos estados comportamentais do agente.

Finalmente, para o enunciado (3):

- **R**₃ = eventos relacionados à disponibilidade do objeto no local da ação;
- **R'**₃ = eventos relacionados às preferências do agente a um determinado tipo de objeto.

Dessa forma, podemos estabelecer a seguinte forma geral para uma questão-por-quê:

Por que **(P)** em contraste com **(X)** (dada uma determinada relação de relevância **(R)**)?

Ademais, em termos formais, podemos também definir uma questão-por-quê como uma tríade ordenada: **Q=<P, X, R>**.

Direcionemos agora a nossa atenção para a explicação propriamente dita, ou melhor, analisemos mais de perto a resposta a uma determinada questão-por-quê **(Q)**. Para tanto, suponhamos que estamos interessados nas razões que levaram Adão e não Eva, por exemplo, a comer a maçã, situação que pode ser figurada, conforme os parâmetros apresentados acima, pelo tema da questão **(P₁)** e a classe de contraste **(X₁)**. Além disso, vamos considerar ainda que estamos interessados em fatores explicativos que levem em consideração a proximidade de Adão da macieira, fato que pode ser figurado pela relação de relevância explicativa **(R₁)**.

Dessa forma, a questão: (1) Por que **Adão** comeu a maçã? Poderia ser representada pela tríade ordenada, **Q = <P₁, X₁, R₁>**, ou ainda de modo mais explícito por:

Por que **Adão** comeu a maçã em contraste com **{Adão, Eva, Caim, Abel}** (dada a seguinte relação de relevância — “**eventos relacionados à disposição espacial do agente em relação ao objeto**”)?

Por conseguinte, uma possível resposta, ou melhor, uma possível explicação poderia ser:

(I) Foi Adão que comeu a maçã, e não outra pessoa, porque **apenas ele estava próximo da macieira**.

Por seu turno, a questão: (2) Por que Adão **comeu** a maçã? Poderia ser representada pela tríade ordenada, **Q = <P₂, X₂, R₂>**, ou ainda de modo mais explícito por:

Por que Adão **comeu** a maçã em contraste com {**comeu, ofereceu, arremessou**} (dada a seguinte relação de relevância — “**eventos relacionados aos estados fisiológicos do agente**”)?

Nesse caso, poderíamos apresentar a seguinte explicação:

(II) Adão comeu a maçã, em vez de oferecê-la a Eva, porque **estava com fome**.

Finalmente, a questão: (3) Por que Adão comeu a **maçã**? Poderia ser representada pela tríade ordenada, $Q = \langle P_3, X_3, R_3 \rangle$, ou ainda de modo mais explícito por:

Por que Adão comeu a **maçã** em contraste com {**maçã, morango, banana**} (dada a seguinte relação de relevância — “**eventos relacionados à disponibilidade do objeto no local da ação**”)?

Finalmente, nesse caso, uma possível explicação poderia ser:

(III) Foi a maçã que Adão comeu porque **não havia outras frutas no Jardim do Éden**.

Portanto, tendo em mente a estrutura das respostas (I), (II) e (III) acima, pode-se dizer que uma explicação apresenta a seguinte forma geral:

(P) em contraste com (o restante de) **(X)** porque **(A)**. (Em que **(A)** é um enunciado qualquer que está numa relação de relevância explicativa **(R)** com o tema da questão **(P)** e a classe de contraste **(X)**).

Ao final da sua proposta, van Fraassen apresenta as condições que tornariam uma explicação **(A)** uma resposta adequada a uma determinada questão-por-quê. Com base em um conjunto específico de regras, o autor avalia quais seriam essas condições. Por exemplo: se (i) a explicação **(A)** é provavelmente verdadeira e (ii) considerando que **(A)** pode explicar outros membros da classe de contraste **(X)**, até que ponto podemos afirmar que tal explicação é mais favorável ao tópico da questão

(P)? Além disso, (iii) se existem ou não outras explicações que seriam mais favoráveis ao tópico (P) do que a explicação (A).

5.1 Um exemplo de uma explicação pragmática

Para um melhor entendimento e detalhamento da formulação de van Fraassen, examinemos a seguinte situação. Suponhamos que um membro de um grupo de engenheiros de um órgão regulador, durante uma visita técnica a uma subestação elétrica, ao observar um determinado condutor deformado, faz a seguinte pergunta ao chefe técnico local que está conduzindo os visitantes naquele momento:

Por que aquele condutor está deformado?

O questionador presume que o condutor está deformado, e está pedindo uma razão para isso. Chamemos a proposição de que o condutor está deformado de o *tema* da questão [...]. Em seguida, essa questão possui uma *classe de contraste*, como vimos, isto é, um conjunto de alternativas. Vou supor que essa classe de contraste, chamemo-la X, é a classe das proposições que incluem o tema. Para essa interrogação em particular, o contraste poderia ser que é esse condutor em vez *daquele*, ou que esse condutor se deformou em vez de manter a sua forma. (VAN FRAASSEN, 2007, p.250–251)

Ao iniciarmos a modelagem de tal situação, podemos supor, como tema da questão, **P** = "aquele condutor está deformado"; e as seguintes classes de contraste (**X**):

- **X₁** = {P, P'}; onde: **P** = "aquele condutor está deformado" e **P'** = "aquele outro condutor mantém a forma";
- **X₂** = {P, P''}; onde: **P** = "aquele condutor está deformado" e **P''** = "aquele (mesmo) condutor mantém a forma".³³

³³ A classe de contraste (**X₁**) representa a situação na qual o questionador deseja entender por que, entre dois condutores distintos, apenas um deles está deformado. Por outro lado, a classe de contraste (**X₂**) ilustra a situação em que o questionador quer saber por que, dentre dois possíveis estados do mesmo condutor (deformado ou não), este se encontra no estado deformado.

Finalmente, há aquilo a cujo respeito uma razão é solicitada, que determina o que vai contar como um possível fator explicativo, a relação de *relevância explicativa*. No primeiro exemplo, a solicitação poderia ser *por eventos “que conduzissem” à deformação*. Isso admite como algo relevante um relato de erros humanos, de interruptores sendo ligados ou a umidade do ar se condensando nesses interruptores, e mesmo encantamentos proferidos por bruxas (uma vez que a avaliação do que é uma resposta adequada vem depois). Por outro lado, os eventos que conduzem à deformação poderiam ser bem conhecidos, caso no qual é provável que a solicitação deva ser pelas condições permanentes que tornaram possível que aqueles eventos levassem a essa deformação: a presença de um campo magnético de certa força, digamos. (VAN FRAASSEN, 2007, p.251)

Dessa forma, por exemplo, quanto à relação de relevância explicativa (**R**), podemos considerar as seguintes possibilidades:

- **R₁** = eventos relacionados a erros humanos na operação da subestação;
- **R₂** = eventos relacionados à atuação automática de determinados interruptores;
- **R₃** = eventos relacionados à umidade do ar se condensando em determinados interruptores;
- **R₄** = eventos relacionados a encantamentos de bruxas;
- **R₅** = eventos relacionados à presença de um campo magnético;

Conforme mencionamos acima, podemos definir uma questão-por-quê como uma tríade ordenada: **Q=<P, X, R>**. Assim, nesse caso, podemos ter o seguinte conjunto de possíveis questões-por-quê:

$Q_1 = \langle P, X_1, R_1 \rangle$	$Q_2 = \langle P, X_1, R_2 \rangle$	$Q_3 = \langle P, X_1, R_3 \rangle$	$Q_4 = \langle P, X_1, R_4 \rangle$	$Q_5 = \langle P, X_1, R_5 \rangle$
$Q_6 = \langle P, X_2, R_1 \rangle$	$Q_7 = \langle P, X_2, R_2 \rangle$	$Q_8 = \langle P, X_2, R_3 \rangle$	$Q_9 = \langle P, X_2, R_4 \rangle$	$Q_{10} = \langle P, X_2, R_5 \rangle$

Por exemplo, de modo formal, ao buscar por razões que expliquem a deformação daquele condutor específico e a manutenção da forma daquele outro, considerando a possibilidade da ocorrência de erros humanos na operação dos

equipamentos da subestação, um questionador pode enunciar a seguinte questão-por-quê:

Q₁: Por que aquele condutor deformou, dado como classe de contraste $X_1 = \{\text{"aquele condutor está deformado"}, \text{"aquele outro condutor mantém a forma"}\}$, e como base explicativa (**R₁**) eventos relacionados a erros humanos na operação da subestação?

Ou ainda, de modo formal, se estiver buscando por razões que expliquem a deformação daquele condutor específico em vez da manutenção de sua forma, considerando a ocorrência de possíveis fenômenos eletromagnéticos, o questionador pode enunciar esta outra questão-por-quê:

Q₁₀: Por que aquele condutor deformou, dado como classe de contraste $X_2 = \{\text{"aquele condutor está deformado"}, \text{"aquele (mesmo) condutor mantém a forma"}\}$, e como base explicativa (**R₅**) eventos relacionados à presença de um campo magnético?

Vale notar que em situações reais de fala, em vez de proferir de modo formal a questão **Q₁** ou a questão **Q₁₀**, um questionador na maioria das vezes enuncia apenas — **Q**: Por que aquele condutor está deformado?

5.2 Os pressupostos de uma questão-por-quê

Por seu turno, no intuito de responder corretamente à pergunta (**Q**), um possível explicador deve considerar um fato essencial que acompanha o enunciado de qualquer questão, ou seja, toda questão possui pressupostos:

Numa pergunta, o pressuposto é uma proposição supostamente aceitável para o oponente, que ficará comprometido com ela ao dar qualquer resposta direta. Em geral, uma pergunta pode ter vários pressupostos. Como um pressuposto é também uma proposição, o ato de fazer uma pergunta contém informações positivas em forma de proposições. Por isso, perguntar pode ser um modo de introduzir proposições num diálogo. Ou seja, perguntar é também argumentar e influenciar o curso subsequente da argumentação. (WALTON, 2012, p.39)

Por exemplo, para um explicador inferir que um questionador, ao enunciar **(Q)**, está precisamente querendo dizer **(Q₁)** ou **(Q₁₀)**, o contexto da interação linguística deve, por assim dizer, preencher as lacunas. No sentido de analisar esse aspecto, retornemos ao exemplo da subestação, tendo agora em mente o seguinte conjunto de circunstâncias. O questionador é um fiscal de um órgão regulador. Seu grupo está em uma visita de rotina. O grupo acabou de verificar que todos os equipamentos e controles automáticos da subestação estão funcionando adequadamente. Em um dado momento, ao observar a deformação inusitada de um condutor da subestação e a manutenção da forma de outro, o fiscal solicita os manuais de operação e manobra dos equipamentos da subestação ao chefe técnico que acompanha o grupo. Após detida leitura, num gesto ostensivo, ao apontar para um condutor (que está deformado) e depois para outro (que não está), ele enuncia a questão-por-quê **(Q)** — “Por que aquele condutor está deformado?”.

Dadas essas circunstâncias, ao responder **(Q)**, o possível explicador está considerando que a questão-por-quê se refere a uma determinada circunstância factual (a deformação daquele condutor específico e a manutenção da forma daquele outro) que está sendo observada por todos naquela situação de fala. Além disso, dado um certo conhecimento sobre a importância dos procedimentos técnico-operacionais na operação de uma subestação, bem como a solicitação dos manuais pelo fiscal, todos os envolvidos na interação encontram-se justificados a considerar que uma deformação como aquela pode ter como origem operações indevidas. Assim, naquela situação, desnecessário enunciar explicitamente que se deve considerar a deformação daquele condutor específico em contraste com a manutenção da forma deste outro, ou ainda, de que os possíveis fatores explicativos devam se restringir a procedimentos operacionais que não foram executados segundo os manuais. Portanto, dado um certo universo de possibilidades, não parece haver dúvida que o questionador está se referindo a uma questão-por-quê formal bem específica, isto é - **(Q₁)** -, ao enunciar **(Q)**.

Formulando de modo mais preciso, uma questão-por-quê ao ser enunciada, apresenta certos balizadores, ou melhor, certos pressupostos que uma vez corretamente inferidos pelo explicador acabam por guiar a posterior explicação. De acordo com a sua natureza, **podemos categorizar esses pressupostos em semânticos, pragmáticos propriamente ditos e comunicacionais.**

Os **pressupostos semânticos** dizem respeito ao fato de que, se estamos envolvidos em uma determinada interação explicativa, os elementos **(P)**, **(X)** e **(R)**, que embasam a questão-por-quê ($Q = \langle P, X, R \rangle$), devem preencher certos requisitos, ou, de modo mais preciso, conforme assevera van Fraassen (2007):

- a) o tema da questão **(P)** deve ser verdadeiro;
- b) somente **(P)**, entre os elementos da classe de contraste **(X)**, deve ser verdadeiro;
- c) há pelo menos uma explicação **(A)**, também verdadeira, que está em uma relação de relevância **(R)** com o par ordenado composto pelo tema da questão **(P)** e a classe de contraste **(X)**.

Por exemplo, dada a questão-por-quê, “por que o vidro daquela janela está quebrado?”, o tema da questão **(P)** = “o vidro daquela janela está quebrado” deve ser verdadeiro. Assim como, ao considerarmos a classe de contraste **(X)** = {"o vidro daquela janela está quebrado", "o vidro daquela janela está intacto"}, apenas o tema da questão **(P)** = “o vidro daquela janela está quebrado” deve ser verdadeiro, isto é, a outra opção, “o vidro daquela janela está intacto”, não se concretizou.

Ademais, o questionador indica a natureza dos eventos que contam como fator explicativo. Em outras palavras, ele explora o contexto ou mesmo executa alguma ação mais explícita no sentido de sugerir a natureza da relação de relevância **(R)**. Assim, ao praticamente escolher a relação de relevância, ele está, por assim dizer, delimitando uma nuvem de possibilidades explicativas de certa natureza, dentro da qual acredita existir uma explicação **(A)** verdadeira. Vale notar que a explicação **(A)** seria apenas um evento dos muitos de mesma natureza que estariam numa relação de relevância **(R)** com o par **(P)** e **(X)**. No caso da janela quebrada, a relação de relevância explicativa **(R)** poderia ser, por exemplo, “atos ostensivos de vandalismo”, e uma possível explicação **(A)**, “uma pedra quebrou o vidro”.

Por seu turno, os **pressupostos pragmáticos propriamente ditos** materializam-se na escolha dos elementos **(P)**, **(X)** e **(R)** segundo as intenções, os desejos e as crenças do questionador. Vale ressaltar que essa escolha, além de procurar induzir uma resposta, busca orientá-la em um sentido específico, não deixando também de considerar a capacidade do explicador em fornecê-la naquelas

circunstâncias. Tais pressupostos indicam; portanto, que uma questão-por-quê não é neutra; sua enunciação reflete de alguma forma os interesses e as crenças do questionador. No exemplo da subestação, a escolha do tema **(P)** e da classe de contraste **(X)** não é aleatória, mas baseada nas intenções e crenças do questionador (o fiscal) sobre o que é importante naquele contexto. Do mesmo modo, a relação de relevância **(R)** é escolhida com base na crença do questionador de que há uma relação significativa e de determinada natureza (por exemplo, “eventos relacionados a erros humanos na operação da subestação”), na qual se inseriria uma suposta explicação **(A)** a ser apresentada pelo explicador.

Por último, os **pressupostos comunicacionais** relacionam-se ao modo como as informações são transmitidas. Melhor dizendo, o que está sendo comunicado por meio de uma questão-por-quê está sendo transmitido também com o maior grau possível de eficiência comunicativa. Assim, se na enunciação de uma questão-por-quê não estiver explícita a relação de relevância **(R)**, o explicador pode pressupor que o contexto da interação linguística deve estar suprimindo essa ausência de alguma maneira. Por exemplo, durante a visita à subestação, o fato do fiscal solicitar aqueles manuais, já sinalizava aos presentes (e ao possível explicador) a relação de relevância **(R)** que estaria em jogo naquela interação explicativa (“eventos relacionados a erros humanos na operação da subestação”), não havendo a necessidade de explicitá-la na formulação da questão.

Por fim, explorando um pouco mais o exemplo da subestação, uma possível explicação do chefe técnico da subestação poderia ser:

A: um superaquecimento deformou o condutor. Isso aconteceu devido a uma manobra indevida sob carga, quando se operava uma chave seccionadora.

Em suma, podemos observar que o explicador, ao formular a explicação (A), leva em consideração as determinações do questionador. Estas determinações, por sua vez, são figuradas pelos pressupostos da questão-por-quê, como vimos. Grosso modo, a dinâmica funciona da seguinte maneira: quando o questionador formula a questão-por-quê, ele endereça os pressupostos ao explicador. Este, por sua vez, os acolhe, integrando-os às suas

próprias determinações no momento da enunciação da explicação. Contudo, em muitas situações, pode não haver nenhuma explicação que responda à questão-por-quê.

6 A TEORIA DOS ATOS DE FALA E A TEORIA DAS IMPLICATURAS

Diante dos aspectos teóricos apresentados pelo modelo de van Fraassen, podemos agora compreender quais seriam as condições capazes de tornar uma explicação falsa ou falaciosa. Tais condições dependem fundamentalmente dos pressupostos semânticos da questão-por-quê. Isto é, um tema da questão **(P)** pouco evidente ou mesmo inexistente, uma escolha inadequada de uma classe de contraste **(X)**, bem como a adoção de uma relação de relevância **(R)** que aponte para eventos que raramente ou que em nenhuma circunstância conduzem ao tema **(P)**, são suficientes para tornar uma explicação **(A)** falsa ou falaciosa. Entretanto, sabemos que o fato de uma explicação ser falsa ou falaciosa não impede a sua instrumentalização nem os efeitos nocivos advindos de sua expressão. Para avaliarmos melhor essa situação, necessitamos aprofundar nossos estudos. Um quadro mais específico de teorias pragmático-linguísticas pode auxiliar-nos na compreensão dessa situação. Para tanto, inicialmente analisaremos a teoria dos atos de fala de John L. Austin e, na sequência, a teoria das implicaturas de Herbert P. Grice.

A teoria dos atos de fala veio a público por meio de uma série de conferências que foram apresentadas em Harvard no ano de 1955. Posteriormente, as notas das conferências foram publicadas, em 1962, com o título "How To Do Things With Words".³⁴ Nesta obra, Austin inicialmente propõe a ideia de que nossos enunciados podem ser categorizados em duas classes distintas:

- i. a **classe dos constatativos**, na qual os enunciados descrevem eventos do mundo ou algum estado de coisas e, por essa razão, são considerados portadores de verdade;
- ii. a **classe dos performativos**, na qual os enunciados modificam um estado de coisas específico. Tais proferimentos não são classificados como verdadeiros ou falsos, mas sim como bem-sucedidos (felizes) ou malsucedidos (infelizes).

³⁴ Infelizmente, a morte prematura de Austin nos privou da oportunidade de acompanhar a evolução contínua de seu pensamento em primeira mão. Em contrapartida, outros autores, como Searle, assumiram o papel de expandir e diversificar a teoria dos atos de fala.

Estamos diante de um proferimento constatativo, quando, por exemplo, dizemos: "A Olimpíada de 2016, realizada no Rio de Janeiro, apresentou excelentes resultados nas competições de natação". Nesse caso, percebemos facilmente que tal afirmação diz respeito a eventos bem específicos que ocorreram em um determinado período e lugar. Para verificar a sua veracidade, basta uma rápida consulta na internet ou em algum almanaque esportivo.

Por seu turno, quando o presidente do Brasil enuncia durante a cerimônia de abertura da olimpíada, "Declaro abertos os Jogos Olímpicos...", tal proferimento não se refere a nenhuma asserção verdadeira ou falsa. O que se observa, nesse caso específico, é a expressão de um enunciado performativo que pode ser bem-sucedido ou malsucedido. Vale notar também como essa simples enunciação altera a realidade, uma vez que os jogos e as competições são iniciados a partir daquele momento.

Entretanto, o enunciado performativo não está isento de críticas; pode-se muito bem criticá-lo, mas numa dimensão completamente diferente daquela do verdadeiro e do falso. É preciso que o performativo seja emitido numa situação que, em todos os pontos, seja apropriada ao ato em questão: se o autor não está em condições exigidas para agir (e há uma quantidade destas condições), então seu enunciado será, como o nomeamos geralmente, "infeliz" (*unhappy*). (AUSTIN, 1998, p.112)

Ou seja, para que a intervenção na realidade por meio do enunciado performativo tenha sido eficaz, Austin assevera que determinados requisitos devem ser atendidos. Para tanto, o autor apresenta um quadro de requisitos que permite determinar as condições de sucesso de um performativo qualquer. A partir do ajuste do enunciado performativo neste quadro, podemos avaliar se este foi feliz (bem-sucedido) ou infeliz (malsucedido):

- (A.1) Deve existir um procedimento convencionalmente aceito, que apresente um determinado efeito convencional e que inclua o proferimento de certas palavras, por certas pessoas, e em certas circunstâncias; e, além disso, que
- (A.2) as pessoas e circunstâncias particulares, em cada caso, devem ser adequadas ao procedimento específico invocado.
- (B.1) O procedimento tem de ser executado, por todos os participantes, de modo correto e
- (B.2) completo.
- (Γ.1) Nos casos em que, como ocorre com frequência, o procedimento visa às pessoas com seus pensamentos e sentimentos, ou visa à instauração de uma conduta correspondente por parte de alguns dos

participantes, então aquele que participa do procedimento, e o invoca deve de fato ter tais pensamentos ou sentimentos, e os participantes devem ter a intenção de se conduzirem de maneira adequada, e, além disso, (Γ.2) devem realmente conduzir-se dessa maneira subseqüentemente. (AUSTIN, 1990, p.31)

Exemplificando: na cerimônia de abertura dos Jogos Olímpicos, o presidente do Brasil enuncia: "Não tenho nada a declarar". Nesse caso, estaríamos diante de uma violação da condição (A.1). Isto é, embora tanto as pessoas envolvidas quanto o procedimento e o contexto tenham sido adequados, o performativo apropriado que o ato exigia não foi enunciado. Desse modo, o ato foi infeliz (malsucedido), não gerando nenhum resultado (ato nulo). Nesse caso, vale observar que a infelicidade do ato foi óbvia. Contudo, existem situações em que somente muito depois da enunciação é que se percebe tal fato.

Em um notável movimento dialético, Austin percebe, ao longo do desenvolvimento de sua teoria, a insuficiência da divisão entre os enunciados constataivos e os performativos. Por exemplo, em muitos casos, a verdade ou a falsidade dos enunciados constataivos depende da felicidade ou infelicidade de seus elementos constituintes. Isso é observado quando analisamos o seguinte enunciado constataivo: "O presidente do Brasil deu início aos Jogos Olímpicos na noite de 5 de agosto de 2016, ao enunciar — *'Declaro abertos os Jogos Olímpicos...'*". Obviamente, trata-se de um enunciado verdadeiro. Contudo, vale notar que isso é verdadeiro porque o enunciado performativo, "Declaro abertos os Jogos...", foi bem-sucedido (feliz).

Por outro lado, também os enunciados performativos podem ser bem-sucedidos ou malsucedidos numa expectativa em relação às condições de verdade dos elementos que compõem a sua expressão linguística.

Relacionado ao performativo (suponho que o seja) "previno-o de que o touro está por atacá-lo" está o fato, se é este o caso, de que o touro está por atacar meu interlocutor. Se o touro não está por fazer isso, então, sem dúvida, o proferimento "previno-o de que o touro está por atacá-lo" se encontra aberto a críticas, mas não em nenhuma das maneiras que até agora caracterizamos como tipos de infelicidade. Neste caso não diríamos que a advertência foi nula — que quem tentou fazê-la não formulou uma advertência, mas que apenas se utilizou de uma forma de advertência — nem que foi insincera. Sentir-nos-íamos muito mais inclinados a dizer que a advertência foi falsa, ou melhor, equivocada, como pode ocorrer com uma declaração. **Portanto, as**

considerações de felicidade e infelicidade podem infectar as declarações (ou algumas delas) e as considerações de falsidade e verdade podem infectar performativos (ou alguns deles). (AUSTIN, 1990, p.58–59, negrito nosso)

Diante das dificuldades em contrastar os enunciados performativos e os constataivos de maneira precisa, o autor propõe a ideia de que as ações linguísticas devem ser entendidas agora como um ato de fala total. Essa ação linguística complexa resultaria da atuação concorrente de outros três atos, a saber:

- a) **ato locucionário** — trata-se da própria emissão ou expressão do enunciado, sem maiores considerações contextuais ou circunstanciais. Assim, destacado do processo dialógico, o que importa é a correção fonética e gramatical de sua expressão, bem como os seus aspectos semânticos (principalmente suas condições de verdade);
- b) **ato ilocucionário** — no qual se considera a intenção do falante (crenças, desejos etc.) ao intervir, por meio da ação linguística, em um determinado estado de coisas;
- c) **ato perlocucionário** — avaliado a partir do resultado induzido pela ação linguística no comportamento do ouvinte ou da audiência.

A partir dessa nova perspectiva, observa o filósofo:

[A]lém da questão, muito estudada no passado, sobre o que um determinado enunciado significa, há outra questão distinta, que se relaciona com a força, como assim a denominamos, de um enunciado. Podemos ter clareza sobre o que significa "Feche a porta", mas ainda não sabemos muito bem se, proferida em um determinado momento, era uma ordem, uma súplica ou outra coisa qualquer. O que precisamos além da velha doutrina sobre significados é uma nova doutrina sobre todas as forças possíveis de enunciados [...]. (AUSTIN, 1961, p.238)

Portanto, diante dessa nova formulação, o filósofo assevera que numa enunciação qualquer, três aspectos em relação a esse ato de fala total podem ser

apontados: **um significado locucionário, uma força ilocucionária e um efeito perlocucionário.**

[D]istinguimos um conjunto de coisas que fazemos ao dizer algo, que sintetizamos dizendo que realizamos um *ato locucionário*, o que equivale, *grosso modo*, a proferir determinada sentença com determinado sentido e referência, o que, por sua vez, equivale, *grosso modo*, a “significado” no sentido tradicional do termo. Em segundo lugar dissemos que também realizamos *atos ilocucionários* tais como informar, ordenar, prevenir, avisar, comprometer-se, etc., isto é, proferimentos que tem uma certa força (convencional). Em terceiro lugar também podemos realizar *atos perlocucionários*, os quais produzimos porque dizemos algo, tais como convencer, persuadir, impedir ou, mesmo, surpreender ou confundir. (AUSTIN, 1990, p.95)

Tendo em vista esses novos desenvolvimentos, o filósofo reposiciona a sua tese inicial:

Que resta, então, da distinção entre o proferimento performativo e o constativo? Na verdade, podemos dizer que o que tínhamos em mente era isto:

(a) No caso do proferimento constativo, nos abstraímos dos aspectos ilocucionários (e de seus aspectos perlocucionários, também) do ato de fala, e nos concentramos no aspecto locucionário: além do mais, usamos uma noção supersimplificada de correspondência com os fatos — supersimplificada porque ela essencialmente absorve o aspecto ilocucionário. Almejamos alcançar um ideal do que seria acertado dizer em todas as circunstâncias, para quaisquer propósitos, para qualquer tipo de ouvinte, etc. Talvez isto seja algumas vezes realizado.

(b) No caso do proferimento performativo, damos o máximo de atenção a força ilocucionária do proferimento, e nos abstraímos da dimensão da correspondência com os fatos. (AUSTIN, 1990, p.119–120)

Nessa reorganização teórica, as condições de sucesso passam a ter sua aplicabilidade redirecionada para a ideia de ato de fala total. Mais precisamente, as condições de sucesso teriam aplicabilidades específicas para cada um dos três atos constituintes que compõem um ato total, conforme ressalta Friggieri (2014, p.56, tradução nossa):

Portanto, podemos analisar os enunciados a partir de três níveis: locucionário, ilocucionário e perlocucionário; e cada nível é acompanhado por sua própria marca de infelicidade ou falha: falha locucionária (falha sintática e semântica), falha ilocucionária (falha em

garantir a aceitação [ou incapacidade de obter entendimento]) e falha perlocucionária (falha em garantir a conformidade [ou incapacidade de obter adesão ou concordância]).

Desse modo, i) no caso do ato locucionário, as condições de sucesso do enunciado seriam determinadas tanto pelas condições de verdade do conteúdo da enunciação quanto por sua correção sintática; ii) no caso do ato ilocucionário e do ato perlocucionário, as condições de sucesso seriam avaliadas respectivamente segundo a força ou a efetividade de tais atos. Para compreendermos tanto o poder desses novos aspectos teóricos quanto as suas consequências, consideremos o seguinte exemplo ilustrativo.³⁵

Numa manhã de domingo, na banca de jornal do Seu Manoel, no bairro do Tremembé, Alfredo encontra Berenice. Alfredo sempre desejou namorar Berenice. Ele sabe que ela terminou recentemente o seu namoro com Carlos. Sabe também que ela gosta muito de filmes franceses. Além disso, Berenice tem evitado sair de casa nesses últimos tempos. Ela ficou tão deprimida depois do namoro com Carlos que até trancou a matrícula no curso de letras modernas na USP.

(Alfredo)

(1) — Hoje à noite, no Cine Ypê aqui do bairro, o filme “O Maravilhoso Destino de Amelie Poulain” está em cartaz.

(Berenice)

(2) — Tenho prova de Alemão amanhã às sete.

(Alfredo)

— Que pena! Quem sabe numa próxima oportunidade. Até mais!

Dado o nosso conhecimento das circunstâncias que envolvem os interlocutores naquela manhã, analisemos os seus atos de fala. No caso, o enunciado (1) de Alfredo apresentou-se como:

- um sucesso locucionário — o enunciado (1) está sintaticamente correto e o filme em questão está realmente em cartaz no Cine Ypê. Nessa dimensão do

³⁵ Vale observar que estamos adotando na análise do caso ilustrativo, a interpretação sugerida por Friggieri (2014).

ato total, podemos dizer que, segundo a terminologia de Austin, nos termos de seu significado locucionário, o conteúdo do enunciado (1) é uma verdade.

- um sucesso ilocucionário — como um ato ilocucionário, o enunciado (1) apresenta-se como um convite. Essa era a intenção de Alfredo, e a posterior reação de Berenice (a recusa) confirma que o ato foi corretamente entendido nesse sentido. Assim, em termos de sua força ilocucionária, o enunciado (1) foi bem-sucedido.
- um insucesso perlocucionário — como um ato perlocucionário, o enunciado (1) não alcança o objetivo pretendido por Alfredo. Berenice recusou o convite. Desse modo, em termos de seu efeito perlocucionário, o enunciado (1) foi malsucedido.

Por seu turno, o enunciado (2) de Berenice apresentou-se como:

- um insucesso locucionário — embora sintaticamente correto, o conteúdo do enunciado (2) é falso. Berenice não tem nenhuma prova no dia seguinte. Ela está distorcendo deliberadamente a verdade (ela trancou a matrícula no Curso de Letras).
- um sucesso ilocucionário — como um ato ilocucionário, o enunciado (2) expressa-se como uma recusa ao convite. A intenção de Berenice, bem como a posterior desistência de Alfredo, confirmam o sucesso do ato.
- um sucesso perlocucionário — como um ato perlocucionário, o enunciado (2) alcança o objetivo pretendido por Berenice, que é a desistência de Alfredo.

Suponhamos que o filme “O Maravilhoso Destino de Amelie Poulain” saiu de cartaz. “*Star Wars*” é agora o filme da semana. Consideremos, no entanto, que Berenice continua matriculada no curso de letras e que realmente tem prova de alemão na manhã de segunda-feira. Diante dessas novas circunstâncias, parte de

nossa análise muda. O conteúdo do enunciado (1) de Alfredo apresenta-se agora como uma inverdade. Entretanto, não podemos descartar que Alfredo acredita estar dizendo a verdade (ele pode desconhecer, por exemplo, a troca de filmes na semana). Por outro lado, o conteúdo do enunciado (2) de Berenice é agora uma verdade. **Assim, vale notar que a despeito das condições de verdade dos conteúdos dos enunciados (1) e (2), os resultados ilocucionários e perlocucionários permanecem os mesmos.**

De acordo com a nossa particular interpretação, é importante ressaltar a conexão óbvia entre a dimensão ilocucionária e a perlocucionária. Conexão, aliás, inexistente entre estas e a dimensão locucionária de um ato de fala total, como vimos. Ou seja, de modo mais preciso, um ato perlocucionário bem-sucedido é uma condição suficiente, mas não necessária, para a ocorrência de um ato ilocucionário bem-sucedido. Por exemplo, consideremos como ato ilocucionário e perlocucionário respectivamente uma pergunta e uma resposta. Dessa forma, se ocorreu uma resposta, então uma pergunta foi formulada. No entanto, a ausência de uma resposta não implica que uma pergunta não tenha sido formulada. Por seu turno, a ocorrência de um ato ilocucionário bem-sucedido é uma condição necessária, mas não suficiente, para um ato perlocucionário bem-sucedido. Por exemplo, nenhuma resposta pode ocorrer, se uma pergunta não for formulada. Contudo, a formulação de uma pergunta não implica que se receberá uma resposta.

De qualquer forma, tanto a força ilocucionária quanto o efeito perlocucionário não dependem das condições de verdade do significado locucionário. Ou seja, tanto a instrumentalização das ações linguísticas quanto as suas possíveis consequências não dependem das condições de verdade dos elementos que compõem o ato de fala total.

Em relação à formulação desenvolvida por Austin que acabamos de apresentar, Marcondes (2017) salienta que algumas críticas apontam para a pouca ênfase dada pela teoria dos atos de fala ao caráter interacional da linguagem. Ora, embora se possa argumentar que as características contratuais³⁶ dos atos de fala evidenciem esse aspecto interacional, é inegável que o desenvolvimento da teoria se concentra principalmente nas intenções do falante. Além disso, a linguagem na vida

³⁶ De acordo com Austin (1990, p.27), o falante ao enunciar um ato de fala cria um compromisso com o ouvinte: “nossa palavra é nosso penhor”.

cotidiana é muito mais fragmentada, oblíqua e indireta do que supõe essa teoria. Entretanto, isso não invalidaria a sua importância. Ou seja, a teoria dos atos de fala continua a ser uma idealização muito útil, podendo ser complementada por outras propostas mais centradas nos aspectos dialógico-comunicacionais e indiretos da linguagem cotidiana. Nessa linha, a teoria das implicaturas apresenta-se como uma tentativa de preencher essa lacuna teórica.

Apresentada por Herbert P. Grice no artigo "Lógica e Conversação" em 1975, a teoria das implicaturas conversacionais concentra-se principalmente no estudo dos aspectos comunicativos envolvidos em uma interação linguística. De acordo com o autor, o fenômeno comunicativo estabelece-se na expressão e no reconhecimento das intenções dos interlocutores envolvidos no diálogo. O processo é simples: a intenção comunicativa do falante é inferida ou calculada pelo ouvinte com base na análise das circunstâncias apresentadas durante a interação linguística, além da própria expressão literal do enunciado.

No sentido de melhor conceituar esse processo comunicativo, o autor introduz os seguintes neologismos: "*implicitar*" e "*implicatura*".

Suponha que A e B estejam conversando sobre um amigo comum C que está, atualmente, trabalhando num banco. A pergunta a B como C está se dando em seu emprego, e B retruca: "*Oh, muito bem, eu acho; ele gosta de seus colegas e ainda não foi preso*". Neste ponto, A deve procurar o que B estava implicando, o que ele estava sugerindo, ou até mesmo o que ele quis dizer ao dizer que — C ainda não tinha sido preso. A resposta poderia ser algo do tipo "C é o tipo de pessoa que tende a sucumbir às tentações provocadas por sua ocupação", ou "os colegas de C são, na verdade, pessoas muito desagradáveis e desleais", e assim por diante. Naturalmente, será desnecessário A fazer qualquer pergunta a B, pois a resposta, no contexto, é antecipadamente clara. Penso que é claro que tudo o que B implicou, sugeriu, significou etc., neste exemplo é distinto do que B disse, que foi simplesmente que C não tinha sido preso ainda. Quero introduzir, como termo técnico, o verbo *implicitar* (implicate) e os nomes correspondentes *implicatura* (implicature) e *implicado* (implicatum). (GRICE, 1982, p.84)

Assim, em uma interação linguística, de acordo com essa nova terminologia, em todas as ocasiões em que um falante emite um enunciado literal, ele também

procura *implicitar*, ou seja, comunicar, a partir do contexto, o que estaria implícito naquelas palavras, mais precisamente, ele procura deixar subentendida uma “implicatura”. Por seu turno, considerando o contexto da enunciação do falante e a sua expressão literal, o ouvinte busca inferir o que está sendo comunicado, procurando calcular aquela “implicatura”.

Vejamos o seguinte exemplo ilustrativo, que nos auxiliará em um melhor entendimento da terminologia e da formulação griceana. Ao escrever uma carta de recomendação para um aluno que pleiteia uma bolsa de estudos em filosofia, um professor passa a maior parte do tempo discorrendo sobre a excelente caligrafia do pupilo. Obviamente, o aluno terá dificuldades para conseguir a bolsa. Na análise do teor da carta, de acordo com a teoria das implicaturas, haveria o envolvimento de dois significados, a saber:

- o significado literal (o dito) — aquilo que está explicitamente expresso. No caso, o próprio teor da carta, isto é, a descrição dos excelentes dotes caligráficos do aluno;
- o significado do falante (o comunicado) — aquilo que se encontra implícito na ação linguística, ou melhor, a implicatura, em que se depreende a evidente falta de condições do aluno em concorrer à bolsa.

Como mencionamos, os ouvintes procuram a todo momento determinar o que está subentendido, ou melhor, eles procuram a todo instante inferir as implicaturas. Entretanto, não se trata apenas do trabalho interpretativo do ouvinte; o próprio falante tem consciência desse processo e tira proveito disso quando propositadamente visa induzir implicaturas convenientes aos seus propósitos.

Uma rápida digressão. Grice divide ainda as implicaturas em convencionais e conversacionais. As implicaturas convencionais dependem somente da interpretação pelo ouvinte dos elementos lexicais do enunciado. Por exemplo, ao compararmos os seguintes enunciados, (1) “José é rico e generoso” com (2) “José é rico, mas generoso”, ambos do ponto de vista estritamente lógico expressam a mesma proposição, uma vez que possuem as mesmas condições de verdade. Ou seja, (1) e (2) são igualmente enunciados verdadeiros se e somente se José for, ao mesmo

tempo, rico e generoso. Entretanto, em (2), a conjunção adversativa — “mas” — desencadeia ou induz na mente do ouvinte uma implicatura (convencional) de oposição, independentemente do contexto de fala dessa enunciação.

As implicaturas conversacionais, por seu turno, são dependentes das condições presentes na interação linguística. Ou seja, o significado do que o falante comunica pode se alterar dependendo do contexto de fala. Doravante, neste trabalho, quando utilizarmos o termo “implicatura”, estaremos nos referindo à implicatura conversacional.

A teoria griceana, ao procurar explicar e descrever o procedimento envolvido no disparo e na interpretação das implicaturas, baseia-se na ideia de que as trocas linguísticas se processam por meio de um código deontológico interiorizado e tacitamente respeitado pelos interlocutores. Este código se estabelece tendo como fundamento o princípio da cooperação:

Faça sua contribuição conversacional tal como é requerida, no momento em que ocorre, pelo propósito ou direção do intercâmbio conversacional em que você está engajado.

Esse princípio desdobra-se em nove máximas que são agrupadas em quatro categorias, a saber:

- Quantidade — (i) *Faça com que sua contribuição seja tão informativa quanto requerido (para o propósito corrente da conversação);* (ii) *Não faça sua contribuição mais informativa do que é requerido;*
- Qualidade — (iii) *Não diga o que você acredita ser falso;* (iv) *Não diga senão aquilo para que você possa fornecer evidência adequada;*
- Relevância ou Relação — (v) *Seja relevante;*
- Modo — (vi) *Evite obscuridade de expressão;* (vii) *Evite ambiguidades;* (viii) *Seja breve (evite prolixidade desnecessária);* (ix) *Seja ordenado.*

Em relação a tais máximas, vale observar as esclarecedoras analogias apresentadas pelo próprio filósofo:

Como um dos meus princípios declarados é ver a fala como um caso especial ou variedade do comportamento intencional, ou melhor, racional, pode ser interessante notar que as expectativas específicas ou presunções correlacionadas com algumas das máximas precedentes têm seus paralelos na esfera das transações que não são diálogos. Enumerarei rapidamente um desses paralelos para cada categoria conversacional.

1. Quantidade. Se você está me ajudando a consertar um carro, espero que sua contribuição seja nem mais, nem menos do que o exigido; se, por exemplo, num estágio particular eu precisar de quatro parafusos, espero que você me alcance quatro e não dois ou seis parafusos.
2. Qualidade. Espero que sua contribuição seja genuína e não espúria. Se necessito açúcar como um ingrediente para o bolo que você está me ajudando a fazer, espero que você não me alcance o sal; se preciso de uma colher, espero que você não me passe uma colher de borracha usada por prestidigitadores.
3. Relação [Relevância]. Espero que a contribuição seja apropriada às necessidades imediatas de cada estágio da transação; se estou mexendo os ingredientes de um bolo, não espero que me seja alcançado um bom livro ou mesmo uma fôrma (embora esta possa ser uma contribuição apropriada num estágio posterior).
4. Modo. Espero que quem estiver me ajudando deixe clara qual a contribuição que está fazendo e que a execute com razoável rapidez. (GRICE, 1982, p.88–89)

Frequentemente, o princípio da cooperação e suas máximas não são respeitados. Duas formas particulares de desrespeito são importantes: a exploração e a violação das máximas. No caso da exploração, o falante não segue as máximas e está ciente de que o ouvinte tem noção desse fato. Ou seja, o falante sabe que o ouvinte não interpretará literalmente suas palavras, conseguindo calcular, de modo adequado, a implicatura proposta. Este é o caso da pessoa que analisa a carta de recomendação. Naquela situação, ela infere que tanto a máxima da categoria da relevância quanto as máximas da categoria da quantidade não estão sendo propositadamente respeitadas quando deveriam, dado o teor que se espera de um documento daquela espécie. No caso, há um excesso de informação irrelevante sendo usado para descrever a bela caligrafia do aluno. Assim, o ouvinte (o analista), ao perceber que o falante (o professor) não está cumprindo deliberadamente aquelas

máximas, infere que o aluno não possui as qualidades que importam para um candidato a uma bolsa.

Por outro lado, diz-se que ocorre uma violação, quando o falante, ao desrespeitar as máximas, sabe que o ouvinte não tem a mínima noção desse fato. Nessa situação, o falante fornece informações irrelevantes, superficiais ou até mesmo falsas. O ouvinte acredita estar participando de um processo cooperativo, quando, na verdade, não está. Vejamos este exemplo:

Quando um falante viola a máxima da quantidade, ele não fornece ao ouvinte informações suficientes que permitam o conhecimento sobre o que está sendo falado, uma vez que ele não quer que o ouvinte conheça o quadro completo. O falante não deseja implicar nada; ele está, por assim dizer, "sendo econômico com a verdade". Você deve conhecer o filme de Peter Sellers, *A Pantera Cor-de-Rosa*. Em determinado momento, seu personagem pergunta a uma recepcionista de um hotel sobre um cachorrinho ao lado da mesa:

A: Seu cachorro morde?

B: Não.

A: [abaixa-se para acariciá-lo e é mordido] Ai! Você disse que o seu cachorro não mordia!

B: Mas este não é o meu cachorro.

A recepcionista sabia que ele estava se referindo ao cachorro à sua frente e não ao seu próprio cão que se encontra na sua casa, mas ela intencionalmente não fornece a ele todas as informações suficientes, por razões somente conhecidas por ela. (CUTTING; FORDYCE, 2021, p.29, tradução nossa)

A exploração e a violação das máximas são bons exemplos de que não existe relação entre as condições de verdade do significado literal e as do significado do falante, conforme assevera Grice:

Já que a verdade de um implicado conversacional [significado do falante] não é requerida pela verdade do que é dito [significado literal] (o que é dito pode ser verdadeiro — o que é implicado pode ser falso), a implicatura não é veiculada pelo que é dito, mas somente pelo dizer o que é dito, ou por “expressá-lo deste modo”. (GRICE, 1982, p.103)

Searle (1999) fez uma análise metódica desses aspectos semânticos, distinguindo **a intenção de comunicar algo** da **intenção de expressar algo de maneira significativa**, ou melhor, de dizer algo segundo as condições apropriadas de verdade.

A intenção de falar de maneira significativa por meio de palavras não deve ser confundida com a intenção de comunicar esse significado a um ouvinte. Normalmente, o principal objetivo de falar é comunicar-se com um ouvinte, mas a intenção de comunicar não é idêntica à intenção de significado, que é a intenção de que um enunciado tenha condições de verdade ou outras condições de satisfação. Então, o que é a intenção de comunicar? [...] Grice percebeu corretamente que, **quando nos comunicamos com as pessoas, somos bem-sucedidos em produzir nelas um entendimento, ao fazê-las reconhecer a nossa intenção de produzir esse entendimento.** A comunicação é singular entre as ações humanas, na medida em que somos bem-sucedidos ao produzir um efeito pretendido no ouvinte, ao fazê-lo reconhecer a intenção de produzir esse mesmo efeito. [...] Como isso funciona? Quando pretendo me comunicar, eu tenho a intenção de produzir entendimento. No entanto, o entendimento consiste em compreender o meu significado. Portanto, a intenção de comunicar é a intenção de que o ouvinte reconheça o meu significado, isto é, que ele me entenda. (SEARLE, 1999, p.144–145, tradução nossa, negrito nosso)

Assim, quando digo, “Es regnet” (Está chovendo! Em alemão), tenho **a intenção de significar algo** com essas palavras. No entanto, ao pronunciar, “Es regnet”, tenho também **a intenção de comunicar algo**, ou seja, tenho **a intenção de que o ouvinte reconheça minhas intenções de significar algo** com essas palavras, ou melhor, **que ele me entenda.**

Assim, seguindo os passos da enunciação “Es regnet”, o meu significado enquanto falante e a minha intenção de comunicação se configurariam da seguinte maneira: Eu enuncio a frase “Es regnet” com a intenção de que :

1. Eu esteja enunciando corretamente uma sentença em alemão com seu significado convencional;
2. Meu enunciado tenha condições de satisfação, notadamente, a condição de verdade de que está chovendo; e
3. O ouvinte reconheça a intenção 2, e ele deve reconhecê-la por meio do reconhecimento da intenção 1 e do seu conhecimento das convenções da língua alemã.

Se o ouvinte reconhecer as intenções 1 e 2, terei sucesso em alcançar a intenção 3. Ou seja, se o ouvinte conhece o idioma, reconhece minha intenção de produzir uma sentença nesse idioma e reconhece que não estou apenas enunciando a sentença, mas também que quero dizer aquilo, então, terei sucesso em comunicar ao ouvinte que está chovendo. (SEARLE, 1999, p.145–146, tradução nossa)

No entanto, do ponto de vista semântico, aquilo que é visado pela intenção de comunicar e aquilo visado pela intenção de enunciar algo significativamente não mantêm uma relação entre si:

Vale notar que essa análise é independente da questão se estou dizendo a verdade ou mentindo, ou se estou sendo sincero ou insincero. Conseguirei fazer a afirmação de que está chovendo mesmo que esteja mentindo. Este é um ponto crucial: mesmo que eu esteja mentindo, ao dizer algo e querer dar significado ao que digo, me comprometo com a verdade do que digo. Portanto, posso ter um compromisso com a verdade mesmo se, de fato, acredito que o que estou dizendo é falso. (SEARLE, 1999, p.146, tradução nossa)

Se o ouvinte reconhecer a intenção do falante de significar algo com as palavras que enuncia, então a intenção de comunicação do falante estará satisfeita. No entanto, isso não impede que o conteúdo do que o falante diz seja falso. Portanto, quando este enuncia algo, os aspectos comunicativos desencadeados pelo seu ato não dependem das condições de verdade do que foi explicitamente dito por ele.

6.1 Como um discurso explicativo comprometido pode ser tão efetivo

Nosso objetivo, ao introduzir a teoria dos atos de fala e a teoria das implicaturas, conforme mencionado anteriormente, era obter, através do estudo de suas bases teóricas, uma visão mais detalhada do caráter linguístico-pragmático de um discurso explicativo falso ou falacioso. Dessa forma, de acordo com a formulação de Austin, tendo em vista o que foi apresentado até este ponto, fica evidente que tanto a questão-por-quê quanto a explicação podem ser pensadas como atos de fala total. Por sua vez, a proposta de Grice também se mostra relevante para um melhor entendimento dos aspectos comunicacionais envolvidos nesses atos.

Como dissemos, quando uma questão-por-quê é enunciada por um questionador, além de sua expressão literal, os pressupostos inerentes a sua enunciação são igualmente endereçados ao explicador. Este, ao recepcioná-los, acaba por incorporá-los na enunciação da explicação. Dessa forma, pode-se dizer que a enunciação de uma explicação, de certo modo, expressa também os pressupostos da questão-por-quê que lhe deu origem.

Para sermos mais específicos, em relação aos pressupostos semânticos da questão-por-quê, a pretensão de verdade que envolve os seus elementos **(P)**, **(X)** e **(R)** é assumida também pelo explicador. Assim, ao enunciar a explicação, o explicador se compromete com a pretensão de verdade desses elementos que somada as próprias condições de verdade do conteúdo da explanação, determinariam a **significação locucionária** desse ato explicativo.

Por outro lado, ao acolher os pressupostos pragmáticos propriamente ditos, o explicador passa a atuar no âmbito factual, teórico e contextual estabelecido pela enunciação da questão-por-quê. Ao proceder dessa maneira, o explicador atribui à enunciação da explicação:

- **uma força ilocucionária** – uma explicação não é neutra, ela procura elucidar a ocorrência do tema **(P)** de acordo com as determinações do explicador. No entanto, essa elucidação está condicionada aos aspectos da realidade que foram recortados e delimitados segundo as intenções e as crenças do questionador, ao determinar **(P)**, **(X)** e **(R)**;
- **um efeito perlocucionário** - com o ato explicativo, o explicador não visa apenas apresentar uma resposta à questão-por-quê, mas também busca a aceitação da audiência ao que está sendo exposto naquele momento.³⁷

Por fim, ao formular uma explicação, o explicador deve também visar à máxima **eficiência comunicativa**. Isso implica transmitir a maior quantidade possível de informação com o mínimo de esforço linguístico do falante e interpretativo do ouvinte. Nesse contexto, de acordo com Grice, é fundamental entender como se dá especialmente a exploração das máximas de quantidade e de modo.³⁸ Essa

³⁷ Este será um ponto a ser tratado no próximo capítulo.

³⁸ A exploração das máximas da quantidade e do modo pode ser ilustrada por meio do seguinte exemplo. Numa determinada situação, André diz:

(1) Maria, abra a porta!

A implicatura que André procura transmitir para a Maria é a seguinte:

(2) Maria, vá até aquela porta, gire a maçaneta no sentido horário até o fim do curso, pare e puxe-a suavemente para você!

exploração das máximas, por exemplo, ocorre quando os elementos ((P), (X) e (R)), possivelmente implícitos na enunciação da questão-por-quê, são mantidos nesse estado no momento do proferimento da explicação. Vale lembrar que, nesses casos, a decisão de tornar tais elementos explícitos ou não (ou melhor, a decisão de recepcionar ou não os pressupostos comunicacionais da questão-por-quê) fica a critério do explicador. Isto é, uma vez que o contexto da enunciação do ato explicativo e o da questão é o mesmo, os elementos anteriormente citados podem permanecer implícitos. No exemplo da subestação, o fiscal ao enunciar – Por que aquele condutor está deformado? – deixa subentendido os elementos (X) e (R). Por sua vez, o explicador também não os menciona no proferimento da explicação, o que não compromete a interação explicativa, como vimos acima.

Neste ponto, vamos finalizar o movimento argumentativo iniciado no **capítulo 4**. Vale lembrar, em forma de asserções e inferências, o caminho percorrido:

- (1) Tanto uma questão-por-quê quanto uma explicação podem ser considerados atos de fala total;
- (2) Uma questão-por-quê apresenta certos balizadores em sua enunciação que podem ser categorizados em **pressupostos semânticos**, **pressupostos pragmáticos propriamente ditos** e **pressupostos comunicacionais**;
- (3) Na enunciação de uma questão-por-quê, o questionador endereça tais pressupostos ao explicador. Este, por sua vez, incorpora-os³⁹ na sua ação explicativa.

Na maioria das situações, expressar (1) já seria o suficiente. Maria entende perfeitamente o que André deseja. O enunciado (1) é conciso, e a sua expressão cumpre o prescrito pelas máximas da Quantidade e pelas máximas de Modo. Entretanto, hipoteticamente, se André enunciar (2)?

Haveria nesse caso um flagrante descumprimento das máximas citadas. O enunciado (2) tem mais informação que o necessário, envolvendo mais esforço tanto no proferimento do falante quanto na interpretação do ouvinte. Somente em situações muito específicas alguém deveria enunciá-lo. Ora, talvez André estivesse apenas querendo fazer graça, inferiria Maria ao ouvir tal enunciado.

³⁹ Mais precisamente, no caso dos pressupostos comunicacionais, vimos que a sua incorporação à explicação fica a critério do explicador.

- (4) Enquanto um ato de fala total, **uma explicação** apresenta um **significado locucionário**, uma **força ilocucionária** e um **efeito perlocucionário**. Além disso, tal ato deve também, ao ser expresso, visar à **máxima eficiência comunicativa**.

Das asserções (1), (2), (3) e (4), bem como dos argumentos desenvolvidos neste capítulo, podemos inferir:

- (5) Os pressupostos semânticos de uma questão-por-quê, que são recepcionados pelo explicador, bem como a veracidade do conteúdo do próprio ato explicativo, determinam o **significado locucionário da explicação**; estabelecendo, nessa dimensão enunciativa, as condições de verdade da ação explicativa.
- (6) Os pressupostos pragmáticos propriamente ditos de uma questão-por-quê, que são recepcionados pelo explicador, estabelecem o campo factual, teórico e contextual em que a explicação se apresentará. O resultado da soma desses fatores com as próprias determinações do explicador estabelece a **força ilocucionária da explicação**. Ademais, ao enunciar a explicação, o explicador visa também à concordância da audiência com o que está sendo proferido ou, em outras palavras, almeja que o seu ato produza um determinado **efeito perlocucionário**. Por fim, no contexto da ocorrência do ato explicativo, tanto a dimensão ilocucionária quanto a perlocucionária são avaliadas segundo condições de sucesso ou insucesso.
- (7) Os **pressupostos comunicacionais** são acolhidos pelo explicador quando os elementos que foram implicitados na enunciação de uma questão-por-quê **((P), (X) ou (R))** são mantidos como tal no proferimento da explicação. Quando isso acontece, podemos dizer que o explicador, ao explorar as máximas griceanas nesse contexto, visa proferir a explicação na **maior eficiência comunicativa possível**.

- (8) O sucesso ou insucesso de um ato de fala total ou, no caso, de uma explicação, em termos da sua **força ilocucionária** ou do seu **efeito perlocucionário**, não depende das condições de verdade da sua **significação locucionária**;
- (9) O que é comunicado e, por conseguinte, a **eficiência comunicativa** de uma explicação, não depende das condições de verdade da **significação locucionária** desse ato de fala total.

Tendo em vista a asserção (5), bem como os fatores que tornam uma explicação falsa ou falaciosa, conforme apresentado no início deste capítulo⁴⁰, conseguimos compreender como o significado locucionário de um ato explicativo pode se tornar uma inverdade. Entretanto, de acordo com as asserções (8) e (9), **isso não impede que essa explicação falsa ou falaciosa, tanto em termos de sua força ilocucionária quanto de seu efeito perlocucionário, seja bem-sucedida, ou mesmo que seja eficiente do ponto de vista comunicativo.**

Dessa forma, compreendemos por que a “razão comprometida”, entendida como um discurso explicativo falso ou falacioso, quando proferida por membros de uma bolha epistêmica, pode ainda ser tão eficaz e persuasiva.

⁴⁰ “[...] um tema da questão (P) pouco evidente ou mesmo inexistente, uma escolha inadequada de uma classe de contraste (X), bem como a adoção de uma relação de relevância (R) que aponte para eventos que raramente ou que em nenhuma circunstância conduzem ao tema (P), são suficientes para tornar uma explicação (A) falsa ou falaciosa.

7 BOLHAS E DESACORDOS EPISTÊMICOS

A interação explicativa inicia-se com a enunciação de uma questão-por-quê. Entretanto, ela não se esgota com o proferimento da explicação. O ato explicativo teria ainda o poder de induzir em quem o ouve ou o lê, as seguintes reações: a aceitação em relação à verdade do que acabou de ser enunciado ou lido, sua recusa ou ainda a suspensão de juízo. A literatura sobre o tema nomeia esses comportamentos de “atitudes doxásticas”. Nesta linha de raciocínio, o comportamento induzido por uma explicação ou qualquer outro elemento que componha um discurso racional, dependeria da compatibilidade entre as atitudes doxásticas dos interlocutores. Assim, caso as partes tenham concordado em relação à verdade de um determinado enunciado, dizemos que as atitudes doxásticas dos envolvidos eram compatíveis naquele momento. Por outro lado, a discordância indicaria a incompatibilidade entre essas atitudes.

Do ponto de vista epistemológico, ao analisarmos as interações racionais em determinados grupos, a maneira como seus membros lidam com possíveis discordâncias pode nos indicar se esses indivíduos estão ou não circunscritos à influência de uma bolha epistêmica. Vale lembrar que uma das principais características de uma bolha é que no seu interior “opiniões discordantes, quando ocorrerem, são sufocadas ou desconsideradas”. Assim, caso observemos tais situações ocorrendo repetidamente dentro de um grupo específico, a probabilidade de estarmos diante de uma bolha epistêmica é bastante elevada. O seguinte exemplo ilustra o que acabamos de dizer.

(M) explica a razão das fortes chuvas extemporâneas que castigam a região em que vive por meio da seguinte enunciação:

p: “a chuva acima da média que foi observada nestas últimas semanas ocorreu devido ao resfriamento das águas do oceano Pacífico. Fato provocado pelo conhecido fenômeno meteorológico *La Niña*.”

Se **(B)** diverge dessa explicação, por acreditar que o grande volume de chuvas se deve a encantamentos de bruxas, dizemos que nesse caso as atitudes doxásticas dos indivíduos são incompatíveis. Ou seja, **(M)** acredita na verdade de “**p**”, enquanto

(B) não. Tudo se complica, se admitirmos por hipótese que tanto (M) quanto (B) são pares epistêmicos⁴¹ e integrantes de um mesmo grupo que compartilham a ideia, entre outras crenças, de que bruxas existem. A despeito da extravagância do exemplo, esse hipotético cenário reflete bem a situação em que contendores como Montaigne (M) e Bodin (B) se encontravam no final do século XVI. Conforme sabemos, Montaigne enfrentou forte constrangimento em seu grupo social ao discordar das interpretações fantasiosas de eventos que poderiam ser explicados de maneira simples e natural. Sua atitude diante das reações provocadas por suas explicações foi o silêncio. Portanto, o sufocamento da explicação divergente do filósofo foi o modo como se resolveu a incompatibilidade entre as suas atitudes doxásticas e as dos membros de seu grupo de pertença naquele momento.

Na análise de um desacordo entre pares epistêmicos, importante ter em mente as seguintes definições. Assim, no exemplo anterior, considera-se uma **evidência de primeira ordem** um determinado conjunto de dados e argumentos que sustentariam a explicação de Montaigne (“p”). Por sua vez, uma **evidência de ordem superior** ou **de segunda ordem** seria aquilo que está sendo *implicitado* quando da expressão de um desacordo. De modo mais preciso, a expressão do desacordo *implicitaria* a opinião de que há um provável problema de confiabilidade nos elementos (dados e argumentos) que comporiam a evidência de primeira ordem.

Estudar as condições da instauração e dos possíveis desenlaces das diversas espécies de desavenças ou desacordos epistêmicos, a despeito destes ocorrerem ou não no interior de uma bolha, é uma das preocupações da nominada epistemologia do desacordo, doravante, ED. Em um dos extremos do espectro de formulações da ED, posicionam-se considerações que em linhas gerais asseveram que a discordância entre dois agentes em pé de igualdade de armas, ou melhor, em paridade epistêmica, enfraqueceria a posição de ambos. O que abriria flancos para uma possível dissolução do desacordo por meio, por exemplo, de uma revisão das atitudes doxásticas dos envolvidos, ou ainda da suspensão de juízo pelos contendores. Conforme a visão de Christensen (2007), por exemplo, a revisão das atitudes pode melhorar os posicionamentos epistêmicos dos envolvidos na desavença. Propostas nesse sentido,

⁴¹ Pares epistêmicos são indivíduos que compartilham um nível equivalente de inteligência, capacidade de raciocínio e domínio sobre o assunto em questão. Além disso, esses indivíduos possuem acesso aos mesmos dados e informações pertinentes para a discussão em pauta.

segundo a literatura sobre o tema, são denominadas de considerações conciliacionistas ou considerações da visão de peso equivalente (*equal weight view*).

De acordo com Christensen, quando estou, por exemplo, em desacordo sobre um determinado assunto com um amigo e há entre nós um reconhecimento mútuo de que estamos em paridade epistêmica, encontro-me diante do seguinte dilema:

Parece-me haver uma pressão considerável para que eu faça pelo menos uma das duas coisas: (1) pensar que meu amigo não deu os pesos relativos certos à simplicidade e ao ajuste com os dados, e não respondeu às evidências com a máxima racionalidade, ou (2) mover minha crença na mesma direção que a dele. (CHRISTENSEN, 2007, p.191, tradução nossa)

Entretanto, segundo a visão de pesos equivalentes, esse dilema não se sustenta. Suponhamos que eu seja um meteorologista bastante experiente e capaz, com um histórico muito consistente de acertos em minhas previsões. Por seu turno, meu colega de profissão é igualmente experiente e capaz e apresenta também um excelente histórico de previsões corretas. Conjecturemos ainda que utilizamos os mesmos dados em nossas previsões, gerados por fontes reconhecidamente confiáveis. Diante de tais informações, a minha previsão que choverá amanhã é de 55%, e a do meu colega é de 45%. Desse modo, qual deveria ser o meu comportamento e o do meu colega em relação a tal desacordo?

Parece-me óbvio que, na ausência de alguma razão especial para pensar que eu tinha alguma vantagem sobre ele ao fazer esta previsão, eu deveria rever a minha crença. Mesmo quando as evidências não proporcionam uma resposta direta à pergunta relevante, a discordância de um par epistêmico fornece razões para a revisão da minha crença. Do meu ponto de vista, isto é positivo: as opiniões das outras pessoas, nestas circunstâncias, apresentam oportunidades para aprimoramento epistêmico. (CHRISTENSEN, 2007, p.194, tradução nossa)

Em suma, na visão de peso equivalente, a expressão da discordância entre os pares epistêmicos ou, para ser mais exato, a evidência de ordem superior que se apresenta a ambos naquele momento, basta para que se inicie a revisão das atitudes doxásticas dos contendores. Ou ainda, em termos técnicos, **a evidência de ordem superior apresenta-se como um derrotador da evidência de primeira ordem nessa visão.**

Por outro lado, ocupando o extremo oposto do espectro de formulações da ED, encontramos as visões não-conciliacionistas. Esse conjunto de considerações advoga simplesmente que pares epistêmicos podem não alcançar um termo em situações de desacordo. Contudo, isso não seria contrário à razão. É o que propõe a nominada visão firme ou resoluto (*steadfast view*). Nessa visão, a evidência de ordem superior que se configura quando da ocorrência de uma divergência, não exige necessariamente nenhuma revisão nos posicionamentos dos pares.

O desacordo não fornece uma boa razão para o ceticismo ou para mudar a visão original de alguém. No que se segue, defenderei a seguinte tese: uma vez que eu tenha examinado minuciosamente as evidências e argumentos disponíveis que sustentam alguma questão, o mero fato de um par epistêmico discordar fortemente de mim sobre como essa questão deve ser respondida, por si só, não tende a minar a racionalidade de se continuar acreditando como eu acredito. Mesmo que eu mantenha com confiança minha visão original diante de tal desacordo, isso não necessariamente constitui uma falha de racionalidade. De fato, manter com confiança minha crença original pode muito bem ser a resposta mais razoável em tais circunstâncias. (KELLY, 2005, p.170, tradução nossa)

Portanto, quando da ocorrência de um desacordo, se determinado contendor acredita firmemente nos elementos que sustentam seu posicionamento, ou de modo mais preciso, se ele crê que a evidência de primeira ordem é robusta naquela situação, a sua atitude de manter-se firme, apresentar-se-ia justificada, a despeito da discordância de seu adversário. **Ou seja, nos termos estabelecidos por Kelly, a evidência de ordem superior nem sempre se apresenta como um derrotador da evidência de primeira ordem.**⁴²

No espectro dos estudos da ED, podemos ainda posicionar, entre as duas perspectivas acima, a denominada visão justificacionista (*justificationist view*). Proposta por Lackey (2010), essa visão aborda exemplos de desacordos incomuns e até mesmo extravagantes. Para ilustrar, a autora apresenta casos extremos de desacordos em que evidências perceptuais aparentemente bem fundamentadas ou até mesmo certezas de natureza matemática são colocadas em dúvida. É o que observamos no exemplo nominado pela autora de “**percepção**”. Vamos imaginar que

⁴² No entanto, de acordo com o autor, admite-se a existência de temas nos quais um possível adversário detém uma posição epistêmica superior. Nessas circunstâncias, se surgir alguma divergência relacionada a esses temas, a evidência de ordem superior pode funcionar como um derrotador da evidência de primeira ordem.

estou almoçando com dois amigos, Estelle e Edwin. Quando peço a Edwin para passar o vinho para Estelle, ele me responde que ela não está presente. No entanto, ela está claramente sentada à mesa, a poucos centímetros de nós. Não estou sob o efeito de álcool ou drogas, nem sinto que esteja alucinando. Ademais, vale notar que nenhum de nós apresentou qualquer problema de ordem cognitiva aparente ou qualquer disfunção perceptiva até o presente momento. Além disso, estamos expressando sinceramente nosso desacordo.

Embora eu tenha tido boas razões para considerar, até agora, Edwin como um par epistêmico, acredito ser racional manter minha convicção de que Estelle está presente à mesa. De fato, mesmo depois de compartilharmos todas as evidências - nas quais Edwin explica que não vê ninguém na cadeira supostamente ocupada por Estelle - ainda me parece racional estar completamente convencido de que ela está almoçando conosco. Dado o grau extraordinariamente alto de confiança justificada que tenho na presença de Estelle, o desacordo de Edwin pode ser visto como a melhor evidência de que algo não está certo com ele, seja em termos de suas considerações evidenciais ou cognitivas. Em outras palavras, parece justificado concluir que Edwin já não é mais um par epistêmico, mesmo que tenha sido antes do desacordo em questão. (LACKEY, 2010, p.307, tradução nossa)

Em outro exemplo, nominado pela autora de “**matemática elementar**”, vamos imaginar agora que eu e meu colega, Harry, discordamos quanto a afirmação de que “ $2 + 2 = 4$ ”. Harry acredita que tal afirmação é falsa:

Como dois adultos com pleno funcionamento das suas faculdades mentais e que possuem evidências igualmente relevantes para o problema em questão, podem discordar sobre se “ $2 + 2 = 4$ ”? Certamente, pelo menos um de nós está confuso ou sofrendo de algum problema cognitivo. Em segundo lugar, quando o desacordo é comum, intuitivamente parece-me bastante racional manter a minha crença, mesmo em face do desacordo com Harry, de quem, aliás, tenho boas razões para acreditar ser um par epistêmico. Pois, dado meu grau extraordinariamente alto de confiança justificada na crença que “ $2 + 2 = 4$ ”, a discordância de Harry pode ser corretamente considerada evidência de que ele não está bem, seja em relação as suas considerações evidenciais ou cognitivas. (LACKEY, 2010, p.308, tradução nossa).

Como vemos, essa visão parte da análise de casos bem específicos nos quais, uma vez estabelecida uma divergência entre dois indivíduos, **(A)** e **(B)**:

- i. **(A)** está altamente justificado a sustentar "p";
- ii. quando se estabelece a divergência, **(A)** está de posse de informação de caráter pessoal de que está em condições cognitivas normais, isto é, não está deprimido, ou sob ação de drogas lícitas ou ilícitas;
- iii. no entanto, **(A)** não tem como avaliar se o outro indivíduo **(B)** encontra-se nessas mesmas condições.

Nessas ocasiões especiais, **tanto (i), (ii) e (iii) quanto a própria natureza extravagante do desacordo (a evidência de ordem superior) transformam-se em evidência de primeira ordem para (A), reforçando e justificando a sua crença de que há algo muito errado com (B). O que, na visão de (A), acaba praticamente por inviabilizar o estatuto de par epistêmico de (B).**

Outro aspecto interessante em nossa discussão é que, a despeito das visões em consideração, não importa se as partes envolvidas em um desacordo epistêmico são grupos *versus* seus integrantes, grupos *versus* um indivíduo qualquer ou grupos *versus* grupos; uma vez que, é sempre possível conceber a interação como se fosse realizada por apenas dois indivíduos. Se um determinado grupo se caracteriza principalmente por ser um conjunto de indivíduos que compartilham as mesmas atitudes doxásticas, então, em termos ideais, podemos reduzir essas atitudes doxásticas grupais às atitudes de um único indivíduo, simplificando a nossa análise. Tal redução é possível devido ao caráter individualista da tradição epistemológica. Tollefsen (2019) menciona os dois principais modos pelos quais o individualismo se apresenta nessa tradição:

A epistemologia tradicional tem sido criticada por apresentar duas formas de individualismo. Em primeiro lugar, considera o indivíduo isolado de seu contexto social. O sujeito "S" que conhece, na expressão "S conhece 'p'", frequentemente encontra-se desvinculado de seu ambiente social, sendo percebido como um conhecedor independente e solitário. Em segundo lugar, a epistemologia tradicional pressupõe que o indivíduo seja o *locus* das propriedades epistêmicas, como crença, conhecimento, cognição, racionalidade,

virtude epistêmica e objetividade (TOLLEFSEN, 2019, p.263, tradução nossa).

Tendo isso em mente, direcionemos a nossa atenção para as ações em que se desconsideram e se sufocam opiniões divergentes no interior de uma bolha epistêmica. Vimos acima, que, além de serem importantes indicadores da ocorrência de uma bolha, tais ações são respostas a um evento específico, no caso, um desacordo epistêmico.

Mais especificamente, vale observar que desconsiderar uma opinião divergente implica, em certo sentido, renegar o estatuto de par epistêmico do indivíduo que diverge da tendência dominante no interior de uma bolha epistêmica. Neste caso em especial, parece-nos adequado supor que esta ação tem como alvo indivíduos que não pertencem a essa bolha. Por outro lado, pode-se observar que sufocar uma opinião discordante não implica nenhuma desconsideração como par epistêmico do indivíduo que diverge. Aliás, a ação preocupa-se mais com o conteúdo da opinião do que com aquele que a enuncia. O que, diante desses aspectos, parece-nos sugerir que o ato de sufocar opiniões aplica-se principalmente a situações de desacordo entre indivíduos pertencentes à mesma bolha.

Diante do exposto, no intuito de se analisar em pormenor essas duas ações, consideremos a ocorrência de um desacordo entre um membro **(R)** de um grupo radicalizado e um indivíduo **(D)**. Suponhamos que o desacordo se dá em torno de uma explicação “**q**” fornecida por **(D)**. Explicação que, por sua vez, contraria fortemente o conjunto de crenças que sustenta a radicalização do grupo ao qual **(R)** pertence.

Primeiro cenário: (D) não é membro do grupo radical ao qual (R) pertence.

Nesse caso, estamos diante de um desacordo intergrupar, e, portanto, dado que os indivíduos não pertencem ao mesmo grupo, a possível ação de **(R)**, ao discordar da explicação “**q**”, é desconsiderá-la. O que equivale a negar o estatuto de par epistêmico de **(D)**, como dissemos.

Ora, segundo a epistemologia tradicional tal situação pode ser reduzida a um desacordo entre dois indivíduos: um agente desafiado **(R)**, cujas atitudes doxásticas refletem as do seu grupo de pertença radicalizado, e um agente desafiante **(D)** que não pertence a esse grupo. Em um desacordo com tais características, a visão

justificacionista de Lackey permite-nos compreender a atitude do agente radicalizado **(R)**. Ou seja, nessa situação, uma vez que **(R)** se considera em pleno controle de suas capacidades cognitivas, podemos supor que, de acordo com a perspectiva radical de seu grupo de pertença, a mera existência do desacordo já seria motivo suficiente para ele inferir que há algo de muito errado com **(D)**, levando-o a negar o estatuto de par epistêmico do desafiante. Neste cenário específico, como se pode observar, quando somente uma das partes é membro de um grupo radicalizado, tanto as prescrições da epistemologia tradicional (o reducionismo a agentes representativos ideais) quanto as considerações da ED (mais especificamente, a visão justificacionista) parecem oferecer uma explicação adequada para as ações do agente desafiado **(R)**.

Segundo cenário: (D) é membro do mesmo grupo radical ao qual (R) pertence.

Por seu turno, em tal cenário, estamos diante de um desacordo intragrupal, e, portanto, dado que os indivíduos agora pertencem ao mesmo grupo, a possível ação de **(R)**, ao discordar da explicação “**q**”, é sufocá-la, como dissemos. Vale salientar que, neste contexto, o agente **(R)** não nega o estatuto de par epistêmico do desafiante. O que ele procura fazer, conforme vimos, é constranger o desafiante **(D)** a abandonar sua posição. Mais precisamente, o constrangimento imposto visa forçar **(D)** a alinhar-se publicamente ao posicionamento do grupo. O desafiante até pode manter suas convicções de forma privada. Contudo, não deve expressá-las publicamente. Neste segundo cenário, não é possível enquadrar teoricamente a situação proposta, seja nas considerações da epistemologia tradicional (onde a redução a agentes ideais se mostra ineficaz, já que ambos os indivíduos pertencem ao mesmo grupo), seja nas considerações da ED. Desse modo, não há meio de se justificar o constrangimento exercido por **(R)**.

Resumindo, diante de um mesmo desacordo, um agente desafiado apresenta diferentes comportamentos a depender da natureza do outro interlocutor. Isto é, o seu comportamento difere caso o agente desafiante seja integrante ou não de seu grupo de pertença. De um modo mais preciso, podemos dizer que tal duplicidade comportamental caracteriza-se por um esforço do agente desafiado em minimizar as diferenças intragrupo e maximizar as diferenças intergrupo.

Priest (2020) propõe um conjunto de ideias que pode nos ajudar a entender tal situação. Para a filósofa, principalmente nos casos em que ocorrem desacordos persistentes,⁴³ as circunstâncias que envolvem o pertencimento a um grupo revestem-se de vital importância:

[A]s forças motivacionais em jogo em muitos desentendimentos de longa data são oneradas pela força motivadora para permanecer leal a um grupo. A motivação da lealdade não está, por necessidade lógica, em conflito com a motivação epistêmica. Entretanto, na prática, os conflitos são frequentes, e as duas, normalmente se reúnem por acidente e não por *design*. Pois, a motivação epistêmica é um fenômeno, e a motivação para mostrar lealdade a um grupo, é outro fenômeno. A finalidade epistêmica é a verdade, a crença justificada, o conhecimento, ou tudo isso. Mas a finalidade da lealdade ao grupo é evitar a traição. (PRIEST, 2020, p.232, tradução nossa)

Ou seja, para a autora, a ocorrência e a persistência de determinadas desavenças ocorrem principalmente em razão da lealdade dos agentes aos compromissos firmados no interior de seus grupos de pertença.

Sempre que surgem evidências contra uma crença em grupo, isto é, quando as evidências sugerem que um grupo possui uma falsa crença, os membros, se desejarem agir com integridade epistêmica, devem apontar os erros epistêmicos ocorridos internamente. No entanto, apontar tais erros significa, frequentemente, abandonar certos compromissos grupais, como, por exemplo, o compromisso de agir como se uma determinada proposição fosse verdadeira, conforme acordado pelo grupo. Nessa situação, temos a motivação epistêmica em conflito com a motivação fundamentada na lealdade ao grupo. Eu defendo que, muitas vezes, a força desta última sobrecarrega a primeira. Quando isso ocorre, é porque os membros do grupo são mais motivados pela lealdade ao grupo do que pelo bem epistêmico e não estão adequadamente motivados epistemologicamente. (PRIEST, 2020, p.232, tradução nossa)

A prevalência dos compromissos de lealdade ao grupo, a despeito das motivações de caráter epistêmico, por exemplo, poderia explicar a razão do constrangimento imposto ao desafiante (**D**) pelo desafiado (**R**), quando da ocorrência de um desacordo intragrupo. Essa ação teria como objetivo restabelecer o compromisso de lealdade. Compromisso que aparentemente se rompeu quando o

⁴³ Podemos pensar que esse conjunto de ideias da filósofa também se aplica àquelas discordâncias que de alguma maneira estariam relacionadas a temas centrais defendidos por grupos polarizados.

integrante desafiante, ao sustentar o desacordo, atacou em público as atitudes doxásticas grupais previamente acordadas.

Assim, podemos dizer que a proposta de Priest, defendida também por Perini-Santos (2020), advoga a tese de que em muitas situações de desacordo envolvendo grupos haveria um conflito entre valores de coordenação, como seriam aqueles relacionados à lealdade ao grupo, e valores epistêmicos. Mais precisamente, **nessas situações, haveria um embate entre uma espécie de normatividade calcada no respeito aos compromissos ou às crenças grupais e a normatividade epistêmica tradicional.**⁴⁴Entretanto, em uma direção inédita, buscando de algum modo superar esse embate, o ramo recente de estudos da epistemologia dos grupos tem procurado definir um novo tipo de normatividade epistêmica. Fundamentada particularmente nas visões não somativas da epistemologia dos grupos, esta nova espécie de normatividade visa trazer para dentro de seu campo de estudos a motivação de lealdade grupal enquanto um compromisso entre indivíduos. Fato que será particularmente útil na compreensão de determinadas situações, como o desacordo intragrupo em uma bolha epistêmica, por exemplo.

⁴⁴ Para sermos mais precisos, os casos que envolvem desacordos entre membros de um mesmo grupo polarizado (desacordos intragrupos) desafiam tanto a normatividade epistêmica tradicional quanto aquela regida pelas formulações da Epistemologia do Desacordo (ED).

8 A EPISTEMOLOGIA DOS GRUPOS

Basta uma rápida consulta na internet para encontrarmos notícias em que desejos, crenças etc. são atribuídas a grupos:

- **Banco Central considera** novo corte da taxa básica de juros: Selic foi reduzida de 3,75% para 3% ao ano na semana passada. E, segundo o BC, deve cair mais em junho. [...] **O Copom justifica**, então, o novo processo de redução da Selic, que pode estimular a demanda e, assim, pressionar a inflação, lembrando que "seu compromisso com a meta para a inflação estabelecida pelo Conselho Monetário Nacional, que sempre foi e segue sendo o principal objetivo da política monetária";⁴⁵
- Luz no fim do túnel: *superávit* primário está mais próximo do que o **mercado imagina**: mesmo com tentativa de furar o teto de gastos, orçamento fictício e novos gastos extraordinários, o país pode ter *superávit* de 0,4% do PIB em 2023;⁴⁶
- Interrupções mexem **com o humor**, mas **mercado acredita** que vacinação irá continuar;⁴⁷
- Como tirar proveito do **comportamento bipolar do Mercado**? O **humor do mercado** às vezes passa do céu para o inferno, não de um dia para o outro. Sabe aquelas pessoas de lua? Aquelas que num dia são as mais felizes e

⁴⁵ Banco Central considera novo corte da taxa básica de juros. **Correio Brasiliense**, Brasília, 12 de maio de 2020. Disponível em: https://www.correiobrasiliense.com.br/app/noticia/economia/2020/05/12/internas_economia,853842/banco-central-considera-novo-corte-da-taxa-basica-de-juros.shtml. Acesso em: 01 de set. de 2021.

⁴⁶ Luz no fim do túnel: *superávit* primário está mais próximo do que o mercado imagina. **Seu Dinheiro**, São Paulo, 02 de jun. de 2021. Disponível em: <https://www.seudinheiro.com/2021/colunistas/seu-dinheiro-convida/luz-no-fim-do-tunel-superavit-primario-esta-mais-proximo-do-que-o-mercado-imagina>. Acesso em: 01 de set. de 2021.

⁴⁷ Interrupções mexem com o humor, mas mercado acredita que vacinação irá continuar. **Valor Econômico**, São Paulo, 25 de fev. de 2021. Disponível em: <https://valor.globo.com/publicacoes/suplementos/noticia/2021/02/25/interruptoes-mexem-com-o-humor-mas-mercado-acredita-que-vacinacao-ira-continuar.ghtml>. Acesso em: 01 de set. de 2021.

simpáticas da face da terra e no dia seguinte são as mais tristes e rabugentas? Pois é, o Mercado Financeiro é pior.⁴⁸

Tais atitudes descrevem estados mentais que dizem respeito a seres que pensam, desejam, creem etc. e que, por esse motivo, são dotados de uma mente. No entanto, é óbvio que grupos não possuem mentes, e embora muitos pensem que ele seja onisciente, onipotente e onipresente, o mercado financeiro não imagina cenários nem apresenta distúrbios emocionais. Dessa forma, a atribuição de estados mentais a grupos não parece gerar maiores consequências em uma análise mais detida, não passando de casos de personificação.⁴⁹ No entanto, em muitas situações, a análise não é tão simples assim. A notícia a seguir diz respeito às apurações da Justiça e do Ministério Público de Minas Gerais sobre os responsáveis pela tragédia de Brumadinho:

A Vale já sabia, desde outubro, que a barragem que rompeu em Brumadinho e outras nove estruturas em Minas Gerais estavam enquadradas em uma categoria de alerta. Essa informação consta de um documento da própria empresa sobre a situação dos reservatórios que a Justiça tornou público nesta terça-feira, 12. Os dados foram obtidos pelo Ministério Público (MP) de Minas, que moveu ação contra a mineradora. A empresa diz que as informações são “probabilidades” e não houve nenhuma detecção prévia de falha nos monitoramentos.⁵⁰

Caso as investigações comprovem que “A Vale já sabia” dos problemas na barragem, de acordo com o nosso ordenamento jurídico, a empresa responderá civil e criminalmente pela tragédia, não eximindo também os seus diretores e técnicos da imputação de crimes. Dessa forma, ir além das considerações metafóricas, visando apreender toda a extensão do significado da expressão, “A Vale já sabia”, é de interesse da sociedade.

⁴⁸ Como tirar proveito do comportamento bipolar do Mercado?. **Leopoldinense**, Leopoldina, 01 de jun. de 2015. Disponível em: <https://leopoldinense.com.br/coluna/371/como-tirar-proveito-do-comportamento-bipolar-do-mercado>. Acesso em: 01 de set. de 2021.

⁴⁹ Figura de linguagem, o mesmo que prosopopeia. Figura pela qual o orador ou escritor empresta sentimentos humanos e palavras a seres inanimados, a animais ou a ideias abstratas.

⁵⁰ Brumadinho: Documentos mostram que Vale sabia de risco elevado de colapso da barragem desde outubro. **Estadão**, São Paulo, 12 de fev. de 2019. Disponível em: <https://www.estadao.com.br/brasil/brumadinho-documentos-mostram-que-vale-sabia-de-risco-elevado-de-colapso-da-barragem-desde-outubro/>. Acesso em: 01 de set. de 2021.

Recentemente, questões relacionadas à atribuição de crenças a grupos são objeto de um novo campo de estudos filosóficos: a epistemologia dos grupos.⁵¹ Para começar, é importante diferenciar o mero agrupamento de pessoas, por exemplo, os alunos que estão de camiseta branca na sala de aula neste momento, de grupos que se encontram razoavelmente estruturados e que exibem, por esse motivo, algum processo de decisão coletivo, como os alunos que apresentarão o seminário na próxima aula. Uma vez estabelecida tal diferenciação, podemos dizer que as formulações filosóficas sobre o tema se dividem em dois grandes ramos: **as considerações somativas e as não somativas.**

As **formulações somativas** afirmam que as crenças de um grupo são simplesmente a soma das crenças de seus membros. Foi o filósofo Anthony Quinton (1975) quem, pela primeira vez, articulou essa visão, inclusive, nomeando-a de “somativa”:

Diz-se que os grupos têm crenças, emoções e atitudes, e que tomam decisões e fazem promessas. Mas essas formas de falar são claramente metafóricas. Ao atribuímos predicados mentais a um grupo, estamos também indiretamente atribuindo-os a seus integrantes. Chamarei de tipo somativo tais atribuições de estados mentais (crenças e atitudes). Desse modo, quando falamos que a classe trabalhadora industrial está determinada a resistir às leis de caráter antissindical, seria o mesmo que dizer que todos ou a maioria dos operários pensa dessa forma. Quando dizemos que grupos decidem ou prometem, estamos nos referindo a declarações institucionais: a referência é a uma pessoa ou pessoas autorizadas a tomar decisões ou assumir compromissos em nome do grupo. (QUINTON, 1975, p.17, tradução nossa).

De acordo com Tollefsen (2015), a visão somativa é adequada às situações em que parece razoável atribuir a mesma crença tanto ao grupo quanto a cada um de seus membros. Nesses casos, diz-se que a crença está distribuída entre os integrantes grupais. É o que acontece quando dizemos que “os torcedores do *Golden*

⁵¹ A epistemologia dos grupos possui fortes vínculos com outro campo filosófico contemporâneo: a ontologia social. Neste ramo de estudo, busca-se responder, por exemplo, a perguntas como: na formação das atitudes doxásticas grupais, estaríamos autorizados a afirmar que existe também uma intencionalidade coletiva sendo formada nessa situação? Se a resposta for afirmativa, os estudos nessa área também procurariam desvendar a origem de tal intencionalidade. A gênese desta estaria no próprio indivíduo ou na coletividade formada pelos integrantes do grupo? Não abordaremos tais temas aqui, pois isso exigiria a discussão de aspectos metafísicos, além de incursões em temas relacionados ao campo da filosofia da mente, o que foge aos nossos propósitos neste estudo.

State Warriors não acreditam que ele disputará as finais da NBA este ano”. De modo razoável, no caso do time se encontrar na penúltima colocação do campeonato, a mesma crença — “de que o time não disputará as finais” — pode ser atribuída indistintamente tanto à torcida quanto a cada um dos torcedores.

A partir dessa perspectiva somativa, podemos afirmar que:

(1) *Um Grupo G acredita que "p" se e somente se todos ou a maioria dos membros de G acreditam em "p".*⁵²

Uma ilustração da aplicação da asserção (1) pode ser observada no **Exemplo diretoria 1**. A diretoria do Banco Central acredita que o PIB do próximo ano será maior que o deste. Assim, segundo a asserção (1), dizemos que todos ou a maioria de seus integrantes também acreditam nesse fato. E, por seu turno, se todos ou a maioria dos integrantes da diretoria acreditam que o PIB será maior no próximo ano; então, a diretoria do Banco Central também deve acreditar nesse fato.

No entanto, suponhamos que apenas o presidente do Banco Central e o seu diretor de política monetária acreditem que o PIB será maior no próximo ano. Nessa situação, vamos supor também que, devido ao poder que possuem na estrutura da organização, as decisões desses agentes determinam praticamente quais devem ser as crenças da diretoria. Assim, ainda se pode afirmar que a diretoria do Banco Central acredita que o PIB será maior no próximo ano, embora a maioria de seus membros individualmente pode não acreditar nisso. Conforme Lackey (2021b), diante da ocorrência de tais casos, faz-se necessária a reformulação da asserção (1), que passa agora a ter a seguinte redação:

(1') *Um Grupo G acredita que "p" se e somente se alguns dos membros de G acreditam em "p".*

Entretanto, mesmo com tal reformulação, a ideia expressa pela asserção (1') não está a salvo de críticas:

⁵² Para reforçar a ideia de que as crenças em questão sejam de conhecimento comum a todos os membros, Tollefsen (2015) adiciona mais uma condição a (1). A nova redação seria a seguinte:

(1) *Um Grupo G acredita que "p" se e somente se todos ou a maioria dos membros de G acreditam em "p", sob condições de conhecimento comum.*

Mesmo com essa modificação, entretanto, alega-se que a abordagem somativa da crença de um grupo sofreria de uma objeção debilitante. Em particular, argumenta-se que um grupo pode adequadamente acreditar que "p", *mesmo quando nenhum de seus membros individualmente acredita em "p"*. Um exemplo clássico disso é quando um grupo aparenta "sustentar" uma opinião grupal, apesar de nenhum dos membros realmente acreditar nela. (LACKEY, 2021b, p.21, itálico da autora, tradução nossa).

O seguinte caso exemplifica a afirmação da filósofa. **Exemplo diretoria 2:** para intermediar as discussões sobre uma nova legislação bancária com a presidência da Câmara dos Deputados, a diretoria do Banco Central indicou como seu representante o funcionário Pedro Figueiredo. Diante desse fato, podemos dizer que, enquanto instituição, a diretoria do Banco Central acredita que Pedro Figueiredo seja a melhor opção para o trabalho. Contudo, nenhum integrante da diretoria acredita realmente que Pedro seja a melhor opção. Individualmente, eles sabem que Igor Freitas é o funcionário mais bem preparado; todavia, indicá-lo causaria muitos problemas com o atual presidente da Câmara — um antigo desafeto de Igor quando este exercia a presidência do Sindicato dos Analistas do Banco Central.

Na consideração desse exemplo, segundo as asserções (1) e (1'), respectivamente, tanto a afirmação — "a maioria dos membros da diretoria do Banco Central acredita..." — quanto a afirmação — "alguns dos membros da diretoria do Banco Central acreditam..." — são falsas, uma vez que nenhum membro de G realmente acredita que Pedro seja a melhor opção. O que implica a falsidade da afirmação — "a diretoria do Banco Central acredita que Pedro seja a melhor opção" — tanto na asserção (1) quanto na asserção (1'). Nessa situação, a crença grupal de que Pedro seja a melhor opção não pode ser distribuída, isto é, não pode ser atribuída a cada membro da diretoria.

Lackey advoga que cenários semelhantes ao **exemplo diretoria 2** podem ter como origem compromissos assumidos no interior do próprio grupo. É o caso, por exemplo, da situação em que duas opções dividem igualmente a opinião grupal. Nessa situação, os integrantes podem concordar em escolher uma terceira opção em que individualmente ninguém acredita. Tais cenários podem também ocorrer quando um regramento externo à ação grupal limita a escolha das opções. É o que acontece quando, devido à instrução de não se considerar testemunhos do tipo "ouvi dizer" ou

“acho que”, um júri inocenta um réu, embora individualmente os jurados acreditem que o réu seja culpado. Há, por fim, cenários que ocorrem por razões pragmáticas. Este é particularmente o caso do **exemplo diretoria 2**, em que questões de ordem pragmática se colocam acima das crenças individuais dos integrantes do grupo.

Os críticos da visão somativa formulam ainda outros cenários extremos nos quais, por exemplo, o fato de todos os integrantes individualmente acreditarem em "p" não implica necessariamente a crença em "p" do grupo. Este é o caso do **exemplo diretoria 3**. Em sua maioria, os integrantes da diretoria do Banco Central acreditam individualmente que o melhor churrasco da cidade é o da Churrascaria Fogo Campeiro. No entanto, enquanto instituição, não se pode dizer que a diretoria do Banco Central acredita que o melhor churrasco da cidade é o da Churrascaria Fogo Campeiro. Ou seja, nesse caso, as crenças distribuídas igualmente entre os membros não podem ser atribuídas ao grupo com um todo.

As **formulações não somativas** foram desenvolvidas tendo em vista principalmente a explicação de cenários semelhantes aos apresentados acima. Este é o caso da visão desenvolvida por Tuomela (1992, 2007). Nesta visão não somativa, o filósofo divide os integrantes grupais em operativos e não operativos. Os membros operativos detêm o poder de decisão no interior do grupo. Por outro lado, os não operativos são destituídos de poder, acatando simplesmente as decisões dos operativos. Para esse caso, podemos formular a seguinte asserção:

(2) Um Grupo G acredita que "p" se e somente se: (i) os membros operativos aceitam conjuntamente que "p" é a visão do grupo; (ii) os membros operativos acreditam mutuamente que (i) é o caso; (iii) em razão da própria estrutura organizacional do grupo definida em (i), caso existam membros não operativos, estes estão inclinados a aceitar tacitamente (ou, no mínimo, deveriam aceitar) que "p"; (iv) o grupo G acredita mutuamente que (iii) é o caso.

Em vista disso, supondo que todos os membros sejam operativos, no **exemplo diretoria 2**, uma vez que todos os integrantes aceitam conjuntamente que a melhor opção para representar os interesses da instituição é a indicação de Pedro Figueiredo; então, podemos dizer que a diretoria do Banco Central acredita nisso também. Por seu turno, supondo ainda que todos os membros sejam operativos, no **exemplo**

diretoria 3, a crença individual dos membros da diretoria de que o melhor churrasco da cidade é o da Churrascaria Fogo Campeiro não se enquadraria nas condições da asserção (2) acima; pois, não se trata de uma crença estabelecida conjuntamente e, por esse motivo, não se pode afirmar que a diretoria do Banco Central acredita nessa afirmação. Por fim, se considerarmos que somente o presidente do Banco Central e o diretor de política monetária são membros operativos e que os demais integrantes da diretoria são membros não operativos; a visão formulada por Tuomela justificaria igualmente a afirmação de que “a diretoria do Banco Central acredita que o PIB será maior no próximo ano se e somente se o presidente do Banco Central e o seu diretor de política monetária acreditarem nisso”.

Uma segunda visão não somativa, que particularmente traz contribuições valiosas ao nosso estudo, foi desenvolvida por Gilbert (1992, 2013). Segundo a autora:

(3) Um grupo G acredita que "p" se e somente se os membros de G aceitarem (ou se comprometerem) conjuntamente que "p".

Isso significa que:

Os membros de G, em conjunto, aceitam que "p" se e somente se for de conhecimento comum em G que seus membros individualmente tenham expressado intencional e abertamente sua disposição em aceitar que "p" junto aos outros membros de G.

Gilbert (1992) esclarece ainda que:

- (1) A aceitação conjunta de uma proposição "p" por um grupo cujos membros são X, Y e Z não implica que haja algum subconjunto do conjunto composto por X, Y e Z, de modo que todos os integrantes de tais subconjuntos também acreditem individualmente que "p";
- (2) Aquele que participa da aceitação conjunta de "p" aceita, portanto, a obrigação de fazer o que for necessário para que, durante a realização de qualquer esforço conjunto entre os membros de G (por exemplo, na ocorrência de reuniões periódicas), sempre se observe como pressuposto a verdade de "p". Além disso, nesse caso, o membro tem o direito de esperar o apoio dos outros integrantes para que isso aconteça;
- (3) Não é necessário aceitar a obrigação de acreditar ou tentar acreditar que "p";

(4) No entanto, se alguém acredita em algo que é inconsistente com "p", é necessário, pelo menos, não expressar essa crença abertamente. Qualquer negação da proposição deve ser qualificada por uma frase que deixe claro que se está falando "*in propria persona*" e não na qualidade de membro do grupo. Esses qualificadores incluem: "Pessoalmente"; "Na minha opinião"; "Falando por mim". (GILBERT, 1992, p. 306–307, tradução nossa)

Posteriormente, a filósofa ressaltará a ideia de que a expressão pública de qualquer crença em desacordo com a crença conjuntamente aceita pelo grupo poderá implicar o constrangimento ou a repreensão do agente divergente pelos demais membros.

Tendo em vista a formulação de Gilbert, como seriam agora analisados os exemplos apresentados acima? No caso do **exemplo diretoria 2**, não importa a crença particular de cada um dos integrantes da diretoria do Banco Central, isto é, a crença de que Igor Freitas seja o funcionário mais bem preparado para a missão. Os integrantes comprometeram-se conjuntamente a indicar Pedro Figueiredo. Por esse motivo, a diretoria do Banco Central, enquanto instituição (como um corpo), acredita que Pedro seja a melhor opção. No **exemplo da diretoria 3**, não importam as preferências gastronômicas da maioria dos integrantes da diretoria. Como não houve o estabelecimento de nenhum compromisso conjunto entre seus membros enquanto instituição, não se pode afirmar que a diretoria do Banco Central acredita que o melhor churrasco da cidade é o da Churrascaria Fogo Campeiro. No que lhe concerne, conforme a visão defendida por Gilbert, também não haveria problema em afirmar que a crença da diretoria foi estabelecida pelo presidente do Banco Central e o seu diretor de política monetária. Se os demais diretores se comprometem a aceitá-la e a defendê-la, a afirmação de que a diretoria igualmente acredita no que foi estabelecido pelo seu presidente e pelo seu diretor encontra-se justificada.

Ao comparar as visões apresentadas acima, Tollefsen (2015) constata:

Relatos somativos parecem explicar atribuições feitas de forma distributiva. Nesse sentido, esses relatos não são sobre crenças grupais, mas sim sobre crenças comuns ou compartilhadas. Entre os relatos que fornecem análises de crenças de um grupo, o de Tuomela parece lidar com o caso de organizações maiores, como corporações. O relato de Gilbert parece ser mais adequado para grupos pequenos e íntimos, como comitês ou mesmo díades. As crenças grupais

parecem ser realizadas de maneiras diferentes por diferentes grupos. (TOLLEFSEN, 2015, p. 36, tradução nossa)

Discordamos apenas da afirmação de que a visão de Gilbert é a mais adequada quando se analisa grupos pequenos. Parece-nos que é mais correto afirmar que essa visão se ajustaria melhor àquelas situações em que temos grupos mais homogêneos e coesos em ação. De qualquer forma, a visão de Gilbert nos oferece subsídios para compreender o respeito ao compromisso com as crenças acordadas em uma bolha epistêmica, bem como as consequências de não ser leal a esse compromisso, conforme mostraremos a seguir.

8.1 A tese da negociação de crença coletiva

No artigo "Conversation and Collective Belief", Gilbert e Priest (2013) estendem o escopo da visão não somativa de Gilbert, analisando também os agrupamentos que se formariam durante um processo conversacional qualquer. Para tanto, as autoras propõem a tese da negociação de crença coletiva. Segundo esta tese, teríamos, de forma paradigmática, a figura de dois ou mais interlocutores participando de uma interação linguística, na qual uma enunciação no mínimo estaria sendo negociada. A partir dessa interação, poderíamos observar a formação de um grupo específico baseado no estabelecimento de um conjunto de crenças compartilhadas. Esse grupo, criado a partir da negociação linguística entre os interlocutores, possuiria as seguintes características: i) seus membros estariam conjuntamente comprometidos com o que foi explícita ou implicitamente expresso nas enunciações acordadas; ii) não seria necessário que as crenças pessoais dos interlocutores concordassem com as crenças estabelecidas por meio das enunciações; iii) no caso de não cumprimento público do que foi conjuntamente acordado, qualquer membro estaria justificado em constranger o transgressor.

Por exemplo, na noite passada, em um curso de filosofia medieval na Faculdade de Filosofia da USP, as provas da existência de Deus desenvolvidas por São Tomás de Aquino foram o tema da aula. O professor detalhou todas as cinco provas propostas pelo Aquinate. Durante a aula, os alunos demonstraram grande interesse, evidenciado pelo intenso debate acadêmico. No final da noite, todos foram unânimes em afirmar a excelência do que foi apresentado e discutido, mesmo que

surpreendentemente muitos fossem ateus. Além disso, em um determinado momento da aula, um aluno ateu expressou publicamente sua contrariedade em relação ao conteúdo da exposição, justificando seu protesto com base no desconforto pessoal provocado pela discussão daquele tema. Esse fato provocou uma forte reação pública de desaprovação ao protesto do aluno por parte dos demais. Não existiram mais protestos depois disso.

Dessa forma, podemos dizer que, durante o período daquela aula, o grupo que se formou ali emergiu a partir de um conjunto de crenças compartilhadas que, por seu turno, foram acordadas por meio de uma negociação pública de declarações. Não é necessário que os participantes possuam crenças semelhantes ou compartilhem algum conhecimento prévio para que a conversação se desenvolva. O importante são as crenças compartilhadas que surgem publicamente no momento da interação linguística. Os proferimentos expressos publicamente, como mencionamos, são objeto de uma negociação que se processa como um tipo de jogo em que alguém propõe algo por meio de um enunciado, e outra pessoa aceita ou recusa aquilo que está sendo expresso (implícita ou explicitamente) nessa ação linguística. Vale observar este outro exemplo de formação pública de crenças, apresentado por Gilbert e Priest:

Dois indivíduos estão sentados um ao lado do outro no metrô. Está muito quente ali. Ambos estão cientes de que está quente e é de conhecimento comum entre eles que está quente. Desse modo, cada um está bem-informado sobre esse fato e sabe que o outro não precisa ser informado sobre isso. Suponhamos também que seja de conhecimento comum que cada um está sofrendo com o calor: dado que ambos estão visivelmente transpirando. Ainda assim, um poderia se dirigir ao outro e dizer: “Nossa, como está quente aqui, não é mesmo?” com o outro interlocutor oferecendo alguma resposta em concordância. Apesar do fato de que ninguém possui novas informações sobre a temperatura naquele instante ou de como o outro está se sentindo, algo importante aconteceu quando eles trocaram essas palavras. Que está quente ali e que cada um está sentindo calor está agora “ao ar livre”, “no espaço público” ou *entre nous*. (GILBERT e PRIEST, 2013, p. 30, tradução nossa)

Embora os compromissos linguísticos se estabeleçam no momento da interação, e um grupo constituído dessa forma possa apresentar um caráter efêmero, nada impede que, diante de encontros frequentes e regulares, ocorra não só a

consolidação de um grupo mais coeso e duradouro, mas também o estabelecimento e a delimitação do campo de ação linguística de seus integrantes.

E se tal interação linguística ocorresse entre indivíduos que pensam de modo semelhante? Conforme estudamos no início deste trabalho, sob certas circunstâncias - como uma ameaça real ou imaginária, um desafio ou a necessidade de resposta a um problema específico - esses indivíduos, ao deliberarem com o objetivo de dar alguma solução para aquela situação premente, acabariam por originar um grupo polarizado ou até mesmo uma bolha epistêmica. Contudo, **vale notar como a gênese desse grupo polarizado ou radicalizado poderia também ser interpretada à luz da proposta de Gilbert e Priest. Ou seja, em relação a formação desse grupo, segundo a formulação das filósofas:**

- sua origem remontaria necessariamente a uma negociação de proferimentos entre indivíduos que, nesse caso particular, pensam de modo semelhante;
- como resultado da negociação linguística na deliberação, seus membros apresentam-se conjuntamente comprometidos com as crenças extremadas formadas a partir dessa interação;
- as crenças pessoais dos membros não precisariam estar alinhadas às crenças extremadas acordadas pelo grupo por meio da interação linguística;
- qualquer membro estaria justificado em constranger qualquer outro integrante que publicamente discordasse do que fora previamente acordado.

A análise da polarização por meio da tese da negociação de crença coletiva de Gilbert e Priest indica-nos, portanto, que discursos proferidos por grupos polarizados carregam necessariamente o estigma de um compromisso em suas enunciações. Compromisso conjuntamente estabelecido durante o processo de deliberação ou de negociação linguística que originou o grupo.

Este seria o motivo de interpretarmos a expressão "razão comprometida" também como razão engajada, razão compromissada; uma vez que discursos⁵³

⁵³ Entre esses discursos, encontra-se o discurso explicativo.

formados no interior dessas bolhas epistêmicas sustentam-se, em última instância, nas crenças conjuntamente acordadas na deliberação ou negociação linguística que deu origem à polarização. Dessa forma, qualquer divergência em relação ao conteúdo de tais discursos implica uma agressão às crenças que sustentam e que originaram aquela bolha. Em outras palavras, qualquer ameaça a esses discursos coloca em risco a coesão e a identidade do próprio grupo. Isso explicaria a pressão a que são submetidos os membros que internamente discordam do pensamento predominante, bem como grande parte do comportamento público aparentemente incoerente, excessivo e até irracional daqueles membros que concordam com tal modo de pensar.

9 CONCLUSÃO

Ao concluirmos nosso trabalho, faz-se necessário tanto fazer um balanço das complexas formulações que utilizamos até o momento quanto apresentar as descobertas decorrentes do estudo. Conforme mencionado anteriormente, as discussões sobre a instrumentalização da razão no interior de grupos polarizados envolvem duas dimensões de significado daquilo que chamamos de "razão comprometida". A cada uma dessas dimensões está associada uma abordagem analítica distinta. Em uma primeira aproximação, o adjetivo "comprometida" é compreendido como a qualidade de algo que sofreu algum dano ou está prejudicado. Toda a análise que segue dessa interpretação concentra-se na resposta à seguinte pergunta:

Por que a enunciação de uma explicação duvidosa ou falaciosa — ou, conforme a terminologia de nosso estudo, uma "razão comprometida" —, ainda que ilegítima, pode ser tão persuasiva no interior de um grupo polarizado persistente?

Conforme definimos na **seção 2.1**, a polarização grupal é um fenômeno que se refere ao processo pelo qual as opiniões e atitudes de agentes grupais que pensam de modo semelhante, após uma determinada deliberação ou discussão, se deslocam para posições mais extremas que seus posicionamentos iniciais. Tendo isso em vista, ao chegar a este ponto em nosso estudo, como vimos nos exemplos apresentados no **capítulo 3**, estamos convencidos de que as consequências das explicações falaciosas no interior de determinados grupos polarizados podem ser devastadoras: inocentes podem ser condenados; reputações podem ser destruídas; vidas podem ser perdidas devido a tratamentos médicos sem embasamento científico; e, por fim, quando parte do que circula na opinião pública são argumentos e explicações falsas ou falaciosas, democracias podem morrer ou mesmo permanecer em um estado agônico.

Frente a tal estado de coisas, a compreensão da interação explicativa como uma atividade humana que se manifesta por meio de atos de fala, bem como a aplicação do modelo explicativo pragmático proposto por van Fraassen, provou-se um

instrumento bastante útil no entendimento dessas situações tão preocupantes, como podemos notadamente constatar da leitura dos **capítulos 4, 5 e 6**.

O modelo de van Fraassen sugere que um discurso explicativo pode ser pensado como um diálogo apropriado entre dois interlocutores, a saber, um questionador propondo uma determinada questão-por-quê ("Por que 'p'?"), e um explicador a respondê-la. Nessa situação, tanto a questão-por-quê quanto a explicação podem também ser pensadas enquanto a expressão de atos de fala específicos. Ainda, segundo o modelo, a própria enunciação da questão-por-quê, por si só, delimitaria o espaço de ação do conhecimento, das informações disponíveis e dos interesses dos envolvidos, determinando efetivamente o modo de expressão da explicação subsequente. Essa delimitação manifesta-se por meio do que chamamos de "pressupostos da questão-por-quê". Esses pressupostos, que são endereçados pelo questionador e posteriormente acolhidos pelo explicador, podem ser categorizados em:

- **pressupostos semânticos**, relacionados aos aspectos veritativos da questão;
- **pressupostos pragmáticos propriamente ditos**, relacionados às atitudes proposicionais do questionador (crenças, intenções etc.) bem como às circunstâncias de fala dos envolvidos na interação explicativa;
- **pressupostos comunicacionais**, relacionados ao fato de que, na enunciação de uma questão-por-quê, o questionador está transmitindo as informações na maior eficiência comunicativa possível.

Ao nos valermos das propostas da teoria dos atos de fala e da teoria das implicaturas em nossa análise, mostramos que a veracidade ou falsidade dos elementos que compõem os pressupostos semânticos de uma questão-por-quê (e que são necessariamente incorporados ao ato de explicar), bem como as condições de verdade do conteúdo da própria explicação, não afeta qualquer possível instrumentalização desta. Em outras palavras, a força e o efeito gerados por uma explicação, bem como a sua eficiência comunicativa, são independentes das condições de verdade de seus elementos constitutivos. Dessa forma, compreendemos por que, ao ser proferida no interior de um grupo polarizado, a "razão

comprometida”, entendida enquanto uma explicação falsa ou falaciosa, se mostra tão eficaz, gerando tantas consequências.

A segunda abordagem analítica em nosso estudo apresenta, como pano de fundo, temas de cunho epistêmico-linguístico. Mais precisamente, esses temas dizem respeito à produção, à sustentação e à transmissão do conhecimento no interior de um grupo polarizado. Diante dessas circunstâncias, a questão a ser respondida agora é:

Por que, mesmo quando fortemente desafiada no interior de um grupo polarizado persistente, a “razão comprometida” é sustentada e defendida com tanta tenacidade por seus integrantes?

A explicação falaciosa resiste e persiste no interior de um grupo polarizado. E tal aspecto somente poderá ser compreendido, se notarmos que existe nessas ocasiões um compromisso firmado entre os seus integrantes para a sua aceitação e defesa. Esse fato sugere que, a partir deste ponto, o adjetivo “comprometida” deve ser interpretado como algo que se apresenta relacionado a um compromisso ou a um acordo entre agentes. Consequentemente, a expressão “razão comprometida” agora deve ser compreendida enquanto razão engajada, compromissada com as crenças conjuntamente acordadas no interior desses grupos. Ademais, amplia-se e enriquece-se a resposta à questão anterior com a caracterização de um grupo polarizado como uma “bolha epistêmica”. Formalmente, conforme definido na **seção 2.3**, um agente encontra-se em uma bolha epistêmica — ou, conforme definimos, em um grupo polarizado persistente — quando é incapaz de distinguir entre o que acredita conhecer e o que efetivamente conhece. Uma “bolha” teria também como características definidoras: a reiteração e o reforço constantes da visão grupal; o sufocamento ou a desconsideração das opiniões discordantes; a delimitação das ações em espaços fechados físicos ou virtuais; e a amplificação dos julgamentos dos principais formadores de opinião.

Ao se observar principalmente o sufocamento das opiniões no interior de uma bolha, presenciamos cenários que desafiam a normatividade epistemológica tradicional. Dessa forma, conforme apresentamos no **capítulo 7**, ao se estabelecer um desacordo entre dois agentes pertencentes à mesma bolha epistêmica, o que se destaca é o constrangimento do agente que desafia o posicionamento grupal. Nessas situações, infere-se a existência de uma motivação não-epistêmica a conduzir essa ação coercitiva. Mostramos que tal motivação apoia-se no sentimento de lealdade ao grupo.

A epistemologia dos grupos, campo de estudos explorado no **capítulo 8**, especialmente na sua vertente não somativa, visa submeter essa motivação de lealdade a considerações de cunho normativo. Dentro dessa visão, a tese da negociação de crença coletiva proposta por Margaret Gilbert e Maura Priest, apresentada na **seção 8.1**, sugere que as crenças grupais emergem de um processo de negociação linguística e de um posterior compromisso com as crenças fundamentadas nesse processo, o que explicaria a tenacidade dos membros em defendê-las publicamente. Portanto, no caso particular de uma bolha epistêmica, essa tese explica o compromisso de seus membros em sustentar e se manter fiéis às crenças grupais mais extremas, bem como o constrangimento ao qual são submetidos aqueles membros que publicamente não se sujeitam a elas.

As ideias apresentadas acima não dependem da natureza das crenças grupais defendidas em uma bolha. Ou seja, uma bolha epistêmica apresenta as mesmas características independentemente daquilo que é defendido no seu interior. Sendo assim, se quisermos agir no sentido de se evitar a formação de tais grupos ou mesmo se quisermos dissolvê-los, parece-nos razoável pensar que qualquer intervenção nesse sentido deve atacar essas características. Em termos gerais, esse tem sido o enfoque de vários estudos recentes.

Por exemplo, os estudos de Levendusky (2018) e de Thomsen, J. e Thomsen, A. (2022) buscaram intervir em grupos polarizados afetivamente⁵⁴, introduzindo informações e perspectivas críticas às posições existentes no interior desses agrupamentos. No estudo de Levendusky, um primeiro conjunto de intervenções visou aumentar a ambivalência partidária dos indivíduos. Isto é, os participantes foram

⁵⁴ Polarização afetiva é aquela que ocorre quando o que motiva a polarização grupal é o sentimento negativo ou hostil em relação aos posicionamentos políticos ou ideológicos do outro grupo.

estimulados a pensar sobre as falhas de seu grupo, ao mesmo tempo, em que refletiam sobre as virtudes do grupo contrário. Um segundo conjunto de intervenções procurou incentivar a independência do indivíduo em relação ao seu grupo, utilizando técnicas que aumentassem o seu sentimento de autoafirmação. O estudo mostrou que essas intervenções foram bem-sucedidas na despolarização de grupos moderadamente polarizados. No entanto, nos grupos fortemente polarizados, as intervenções intensificaram a polarização. No estudo de Thomsen, J. e Thomsen, A., as intervenções buscaram incentivar o diálogo entre grupos. O resultado da intervenção foi positivo entre grupos fracos e moderadamente polarizados. Contudo, nos grupos fortemente polarizados, a intervenção não surtiu efeito.

Em outro estudo, Santoro e Broockman (2022) mostraram que, ao serem incentivados a dialogar com indivíduos de grupos opostos, integrantes de grupos afetivamente polarizados conseguiam chegar a um acordo em questões não relacionadas às suas identidades grupais, resultando em uma diminuição temporária da polarização. No entanto, quando o diálogo abordava assuntos ligados a essas identidades, o consenso não era alcançado, impedindo qualquer efeito despolarizante.⁵⁵

Em suma, os trabalhos acima evidenciam a resistência das bolhas epistêmicas (ou, nos termos dos estudos citados, dos grupos fortemente polarizados) às diversas intervenções propostas. De acordo com o nosso estudo, isso não nos surpreende. A forte coesão grupal sustentada no compromisso entre os membros e a consequente estigmatização daqueles que não pertencem ao grupo explicam os resultados apresentados. Vale notar ainda que, no caso dos grupos virtuais, conforme mostramos, os algoritmos de rede, o controle dos indivíduos sobre as notícias consumidas e principalmente a **homofilia** dificultam a entrada de informações divergentes ao consenso grupal em cenários reais e não controlados.

Portanto, conforme as considerações desenvolvidas neste trabalho, parece que ainda estamos distantes de um conjunto efetivo de políticas e de intervenções que possa fornecer uma resposta satisfatória aos desafios impostos pelo fenômeno da

⁵⁵ Recentemente, uma série de estudos sobre o tema da polarização grupal foi publicada nas revistas *Nature* e *Science*. Mais precisamente, os artigos de Nyhan et al. (2023) e Guess et al. (2023) estudaram a possibilidade de despolarizar determinados grupos (polarizados afetivamente) por meio de intervenções diretas nos algoritmos curadores do *Facebook* e do *Instagram* durante as eleições americanas de 2020. Os estudos concluíram que essas ações foram ineficientes na despolarização dos grupos estudados. Mais uma vez, esses resultados confirmam a resiliência das bolhas epistêmicas.

polarização. Além disso, a recente disseminação de *chatbots*⁵⁶ de inteligência artificial e, por essa razão, a possibilidade de produção massiva de informações falsas, de autenticidade quase inverificável, torna a situação ainda mais complexa.

De qualquer forma, cremos que a pesquisa que desenvolvemos neste trabalho contribui para a compreensão de elementos fundamentais que compõem as intrincadas circunstâncias associadas à formação e à sustentação das bolhas epistêmicas. Apesar da situação atual ser desafiadora e, do nosso ponto de vista, até bastante preocupante, não devemos nos abater. Precisamos encará-la como uma oportunidade de aprofundar e trabalhar as descobertas apresentadas aqui. Nesse aspecto, a negociação linguística que dá origem ao compromisso de lealdade nas bolhas epistêmicas ainda é pouco compreendida e explorada.⁵⁷ Atuar na gênese dessas bolhas, dificultando a sua formação e o seu consequente estabelecimento, apresenta-se, a nosso ver, como um caminho promissor no desenvolvimento de possíveis estratégias de intervenção.

⁵⁶ Sistemas computacionais que se comunicam com humanos, de forma natural e informal, por meio de diálogos.

⁵⁷ Por exemplo, nesse sentido, o estudo desenvolvido por Olsson (2020), que simula, a partir de um modelo bayesiano, como as crenças grupais escalam para a radicalidade quando indivíduos interagem em um grupo, pode ser um interessante ponto de partida.

REFERÊNCIAS

- ANDERSON, Elizabeth. Epistemic bubbles and authoritarian politics. *In*: EDENBERG, Elizabeth; HANNON, Michael (org.). **Political Epistemology**. Oxford: Oxford University Press, 2021. p. 11-30.
- ARISTÓTELES. **Política**. Tradução de Mário da Gama Kury. Brasília: Editora da Universidade de Brasília, 1985.
- AUSTIN, John Langshaw. Performative utterances. *In*: URMSON, J. O.; WARNOCK, G. J. (org.). **Philosophical Papers**. Oxford: Oxford University Press, 1961. p. 220-239.
- AUSTIN, John Langshaw. **Quando dizer é fazer: palavras e ação**. Tradução de Danilo Marcondes de Souza Filho. Porto Alegre: Artes Médicas, 1990.
- AUSTIN, John Langshaw. Performativo-constativo. **Visão Performativa da Linguagem**. Tradução de Paulo Roberto Ottoni. Campinas, p. 111-144, 1998.
- BAKSHY, Eytan; MESSING, Solomon; ADAMIC, Lada A. Exposure to ideologically diverse news and opinion on Facebook. **Science**, [s.l.], v. 348, n. 6239, p. 1130-1132, 5 jun. 2015.
- BALIETTI, Stefano; GETOOR, Lise; GOLDSTEIN, Daniel G.; WATTS, Duncan J. Reducing opinion polarization: effects of exposure to similar people with differing political views. **Proceedings of the National Academy of Sciences**, [s.l.], v. 118, n. 52, p. 1-11, 22 dez. 2021.
- BERTOLOTI, Tommaso; MAGNANI, Lorenzo. An epistemological analysis of gossip and gossip-based knowledge. **Synthese**, [s.l.], v. 191, n. 17, p. 4037-4067, 23 jul. 2014.
- BODIN, Jean. **On the demon-mania of witches**. Tradução de Randy A. Scott. Toronto: Centre for Reformation and Renaissance Studies, 2001.
- CHESHER, Chris. How computer networks became social. **Second International Handbook of Internet Research**. Dordrecht: Springer Netherlands, p. 105-125, 2020.
- CHRISTENSEN, David. Epistemology of disagreement: the good news. **The Philosophical Review**, [s.l.], v. 116, n. 2, p. 187-217, 2007.
- CLARK, Stuart. **Pensando com demônios: a ideia de bruxaria no princípio da Europa moderna**. Tradução de Celso Mauro Paciornik. São Paulo: Edusp, 2006.
- COADY, Cecil Anthony John. Pathologies of testimony. *In*: LACKEY, Jennifer; SOSA, Ernest (org.). **The Epistemology of Testimony**. Oxford: Oxford University Press, p.253-271, 2006.

CUTTING, Joan; FORDYCE, Kenneth. **Pragmatics**: a resource book for students. New York: Routledge, 2021.

DELUMEAU, Jean. **História do medo no ocidente, 1300-1800**: uma cidade sitiada. Tradução de Maria Lucia Machado. São Paulo: Companhia de Bolso, 2009.

DUTRA, Luiz Henrique de Araújo. **Realismo, empirismo e naturalismo**: o naturalismo nas filosofias de Boyd e Van Fraassen. 1993. 298 f. Tese (Doutorado em Filosofia) - Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 1993.

EDWARDS, Ward. The theory of decision making. **Psychological Bulletin**, [s.l.], v. 51, n. 4, p.380-417, 1954.

ELIAS, Norbert; SCOTSON, John. **Os estabelecidos e os outsiders**: sociologia das relações de poder a partir de uma pequena comunidade. Tradução de Vera Ribeiro. Rio de Janeiro: Zahar, 2000.

FAYE, Jan. **The nature of scientific thinking**: on interpretation, explanation, and understanding. London: Palgrave Macmillan, 2014.

FIEDLER, Klaus; JUSLIN, Peter. Taking the interface between mind and environment seriously. *In*: FIEDLER, Klaus; JUSLIN, Peter (org.). **Information sampling and adaptive cognition**. Cambridge: Cambridge University Press, p.3-29, 2006.

FRIGGIERI, Joe. Etiologies. *In*: GARVEY, Brian (org.). **J.L. Austin on language**. New York: Palgrave Macmillan, p. 50-69, 2014.

GILBERT, Margaret. **On social facts**. New Jersey: Princeton University Press, 1992.

GILBERT, Margaret. **Joint commitment**: how we make the social world. Oxford: Oxford University Press, 2013.

GILBERT, Margaret; PRIEST, Maura. Conversation and collective belief. *In*: **Perspectives on Pragmatics and Philosophy**. [s.l.]: Springer International Publishing, p. 1-34, 2013.

GRICE, Paul. Lógica e conversação. DASCAL, Marcelo (org.). **Fundamentos metodológicos da linguística IV**. Campinas: Instituto de Estudos da Linguagem da Unicamp, 1982. p. 81–103.

GUESS, Andrew M. et al. How do social media feed algorithms affect attitudes and behavior in an election campaign? **Science**, v. 381, n. 6656, p. 398-404, 2023.

HALLIDAY, David; WALKER, Jearl; RESNICK, Robert. **Fundamentals of physics**: volume one. 12. ed. Hoboken: Wiley, 2022.

HEMPEL, Carl Gustav. **Aspects of scientific explanation and other essays in the philosophy of science**. New York: The Free Press, 1965.

HUME, David. **Investigações sobre o entendimento humano e sobre os princípios da moral**. São Paulo: UNESP, 2004.

KAHAN, Dan M.; PETERS, Ellen; WITTLIN, Maggie; SLOVIC, Paul; OUELLETTE, Lisa Larrimore; BRAMAN, Donald; MANDEL, Gregory. The polarizing impact of science literacy and numeracy on perceived climate change risks. **Nature Climate Change**, [s.l.], v. 2, n. 10, p. 732-735, 27 maio 2012.

KAHNEMAN, Daniel. A perspective on judgment and choice: mapping bounded rationality. **American Psychologist**, [s.l.], v. 58, n. 9, p. 697-720, 2003.

KAHNEMAN, Daniel. **Rápido e devagar**: duas formas de pensar. Tradução de Cássio de Arantes Leite. São Paulo: Objetiva, 2012.

KELLY, Thomas. The epistemic significance of disagreement. *In*: **Oxford Studies in Epistemology**. Oxford: [s.n.], p. 167-196, 2019.

KOYRÉ, Alexandre. **Estudos de história do pensamento científico**. Tradução de Márcio Ramalho. Rio de Janeiro: Forense Universitária; Universidade de Brasília, 1982.

LACKEY, Jennifer. A justificationist view of disagreement's epistemic significance. *In*: **Social Epistemology**. Oxford: Oxford University Press, p. 298-325, 2010.

LACKEY, Jennifer. Echo chambers, fake news, and social epistemology. *In*: BERNECKER, Sven; FLOWERREE, Amy K.; GRUNDMANN, Thomas (org.). **The Epistemology of Fake News**. Oxford: Oxford University Press, 2021a. p. 206-227.

LACKEY, Jennifer. **The epistemology of groups**. New York: Oxford University Press, 2021b.

LE MENS, Gaël; DENRELL, Jerker. Rational learning and information sampling: on the 'naivety' assumption in sampling explanations of judgment biases. **Psychological Review**, [s.l.], v. 118, n. 2, p. 379-392, 2011.

LEVACK, Brian P. **A caça às bruxas na Europa moderna**. Tradução de Ivo Korytowski. Rio Comprido: Campus, 1988.

LEVENDUSKY, Matthew S. When efforts to depolarize the electorate fail. **Public Opinion Quarterly**, [s.l.], v. 82, n. 3, p. 583-592, 2018.

LYCAN, William G. Explanation and epistemology. *In*: **The Oxford Handbook of Epistemology**. [s.l.]: Oxford University Press, p. 408-433, 14 nov. 2002.

MARCONDES, Danilo. **As armadilhas da linguagem**: significado e ação para além do discurso. Rio de Janeiro: Zahar, 2017.

MESSINGSCHLAGER, Tanja; HOLTZ, Peter. Filter bubbles und echo chambers. *In*: **Die Psychologie des Postfaktischen: über Fake News**, "Lügenpresse", Clickbait & Co. Berlin: Springer Berlin Heidelberg, p. 91-102, 2020.

MIN, Seong Jae. From algorithmic disengagement to algorithmic activism: charting social media users' responses to news filtering algorithms. **Telematics and Informatics**, [s.l.], v. 43, p. 101251, 2019.

MONTAIGNE, Michel de. **Ensaio**. Tradução de Sérgio Milliet. São Paulo: Editora 34, 2016.

MYERS, David G. **Psicologia social**. Tradução de Daniel Bueno, Maria Cristina Monteiro, Roberto Cataldo Costa. Porto Alegre: AMGH, 2014.

MYERS, David G.; TWENGE, Jean M. **Social psychology**. New York: McGraw-Hill Education, 2022.

NGUYEN, C. Thi. Echo chambers and epistemic bubbles. **Episteme**, [s.l.], v. 17, n. 2, p. 141-161, 13 set. 2018.

NYHAN, Brendan et al. Like-minded sources on Facebook are prevalent but not polarizing. **Nature**, v. 620, n. 7972, p. 137-144, 2023.

OLSSON, Erik J. Why Bayesian Agents Polarize. *In*: BRONCANO-BERROCAL, Fernando; CARTER, J. Adam (org.). **The Epistemology of Group Disagreement**. New York: Routledge, 2020.

OSTORERO, Martine. The rise of the witchcraft doctrine. *In*: DILLINGER, Johannes (org.). **The Routledge History of Witchcraft**. New York: Routledge, 2020. p. 61-77.

PARISER, Eli. **O filtro invisível: o que a internet está escondendo de você**. Rio de Janeiro: Zahar, 2012.

PEIRCE, Charles Sanders. Escritos coligidos. *In*: PEIRCE, Charles Sanders; FREGE, Gottlob. **Os Pensadores**. São Paulo: Abril Cultural, 1974. p. 7-192.

PERINI-SANTOS, Ernesto. What is post-truth? A tentative answer with Brazil as a case study. *In*: **Democracy and Brazil**. [s.l.]: Routledge, p. 226-249, 2020.

POPPER, Karl Raymund. **O mito do contexto**. Tradução de Paula Taipas. Lisboa: Edições 70, 1996.

POPPER, Karl Raymund. **Lógica das ciências sociais**. Tradução de Estevão de Rezende Martins, Apio Cláudio Muniz Acquarone Filho, Vilma de Oliveira Moraes e Silva. Rio de Janeiro: Tempo Brasileiro, 2004.

PRIEST, Maura. The mirage of individual disagreement: groups are all that stand between humanity and epistemic excellence. *In*: BRONCANO-BERROCAL, Fernando; CARTER, J. Adam (org.). **The Epistemology of Group Disagreement**. New York: Routledge, 2020. p. 230-258.

QUINTON, Anthony. The presidential address: social objects. **Proceedings of the Aristotelian Society**, London, v. 76, p. 1-27, 1975.

ROBSON, Jon. A social epistemology of aesthetics: belief polarization, echo chambers and aesthetic judgement. **Synthese**, [s.l.], v. 191, n. 11, p. 2513-2528, 20 mar. 2014.

ROSNOW, Ralph L. Rumor and gossip in interpersonal interaction and beyond: a social exchange perspective. *In*: **Behaving Badly: aversive behaviors in interpersonal relationships**. [s.l.]: American Psychological Association, p. 203-232, 2001.

SALMON, Wesley C. **Four decades of scientific explanation**. Pittsburgh: University of Pittsburgh Press, 2006.

SALMON, Wesley C. Scientific explanation. *In*: BIRD, Alexander; LADYMAN, James (org.). **Arguing About Science**. New York: Routledge, 2013. p. 337-365.

SANTORO, Erik; BROOCKMAN, David E. The promise and pitfalls of cross-partisan conversations for reducing affective polarization: evidence from randomized experiments. **Science Advances**, [s.l.], v. 8, n. 25, p. 1-16, 2022.

SEARLE, John Rogers. **Mind, Language and Society**: philosophy in the real world. New York: Basic Books, 1999.

SIMON, Herbert A. Invariants of human behavior. **Annual Review of Psychology**, [s.l.], v. 41, n. 1, p. 1-20, jan. 1990.

SUNSTEIN, Cass R. **A era do radicalismo**: entenda por que as pessoas se tornam extremistas. Tradução de Luciene Scalzo Guimarães. Rio de Janeiro: Elsevier, 2010.

THALER, Richard H. **Misbehaving**: a construção da economia comportamental. Tradução de George Schlesinger. Rio de Janeiro: Intrínseca, 2019.

THOMSEN, Jens Peter Frølund; THOMSEN, Anna Håland. Intergroup contact reduces affective polarization but not among strong party identifiers. **Scandinavian Political Studies**, [s.l.], p. 1-23, 2022.

TOLLEFSEN, Deborah Perron. **Groups as agents**. Cambridge: Polity Press, 2015.

TOLLEFSEN, Deborah Perron. The Epistemology of groups. *In*: FRICKER, Miranda; GRAHAM, Peter J.; HENDERSON, David; PEDERSEN, Nikolaj J.L.L. (org.). **The Routledge Handbook of Social Epistemology**. New York: Routledge, 2019. p. 263-273.

TSEBELIS, George. **Jogos Ocultos**: escolha racional no campo da política comparada. São Paulo: EDUSP, 1998.

TUOMELA, Raimo. Group beliefs. **Synthese**, [s.l.], v. 91, n. 3, p. 285-318, jun. 1992.

TUOMELA, Raimo. **The philosophy of sociality**: the shared point of view. Oxford: Oxford University Press, 2007.

VAN FRASSEN, Bas C. **A imagem científica**. São Paulo: Editora UNESP, 2007.

VOLTAIRE, François-Marie Arouet. **A philosophical dictionary**. Chapel Hill: The Project Gutenberg, 2011. E-book.

WALTON, Douglas N. **Lógica informal**: manual de argumentação crítica. Tradução de Ana Lúcia R. Franco e Carlos A. L. Salum. São Paulo: Martins Fontes, 2012.

WOODS, John Hayden. Epistemic bubbles. *In: **We will show them! Essays in honour of Dov Gabbay***. London: College Publications, 2005. v. 2, p. 731-774.

WOODWARD, James; ROSS, Lauren. Scientific explanation. *In: ZALTA, Edward N. (Ed.). **The Stanford Encyclopedia of Philosophy***. Edição de verão 2021. Disponível em: <https://plato.stanford.edu/archives/sum2021/entries/scientific-explanation/>. Acesso em: 20 de jan. 2021.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância**. Tradução de George Schlesinger. Rio de Janeiro: Intrínseca, 2021.