

Ensaio 10

PROVAR A CAUSA: O HOLISMO DA JUSTIFICAÇÃO E O ATOMISMO DE DAUBERT

A Consiliência de Induções tem lugar quando uma indução obtida a partir de uma classe de fatos coincide com uma indução obtida a partir de uma classe diferente. Essa Consiliência constitui um teste para a verdade da teoria em que ela ocorre.

– William Whewell¹

Como indica o título, este artigo enfoca a prova da causa em litígios provocados por alegações de delitos com substâncias tóxicas; e, como dá a entender o subtítulo, ele propõe dois argumentos principais, o primeiro de cunho epistemológico e o segundo de cunho legal. O argumento epistemológico é que, sob certas condições, um acúmulo de provas justifica a conclusão em grau mais elevado do que qualquer dos seus componentes poderia fazer isoladamente; o argumento legal, interligado com esse fato, é que a lei probatória norte-americana impõe uma espécie de atomismo que pode impedir o processo de chegar à conclusão mais justificada pelas provas – sendo que os efeitos disso têm sido especialmente evidentes no tocante à prova da causa em casos referentes a delitos com substâncias tóxicas.

A seção 1 montará o cenário, examinando alguns casos em que a questão epistemológica a ser abordada aqui mereceu explicitamente a atenção dos tribunais; a seção 2 desenvolverá o argumento epistemológico, primeiro de forma geral e, em seguida, na forma em que se aplica às espécies de prova da causa tipicamente encontradas em litígios provocados por alegações de delitos com substâncias tóxicas; a seção 3 se baseará nessa explanação para responder a algumas das questões epistemológicas a respeito da prova da causa que estiveram

em pauta nesses casos; e a seção 4 desenvolverá o argumento legal, mostrando que, para o cúmulo da ironia, a exigência de Daubert de que o testemunho especializado seja confiável pode às vezes obstruir o caminho de uma apreciação adequada da confiabilidade da prova da causa.

1. Montando o cenário

Mary Virginia Oxendine nasceu em 1971. Seu antebraço direito ficou atrofiado e ela tem apenas três dedos, grudados um no outro, na sua mão direita. Acreditando que os defeitos de nascença de sua filha foram causados pelo fato de sua mãe ter tomado Bendectin contra enjoo matinal enquanto estava grávida de Mary, os Oxendines moveram um processo contra os fabricantes, a Merrell Dow Pharmaceuticals².

No primeiro julgamento, com júri, o dr. Alan Done testemunhou a favor dos Oxendine que certos anti-histamínicos são conhecidos por seus efeitos teratogênicos em animais e que um dos componentes do Bendectin é o anti-histamínico succinato de doxilamina; que estudos com animais realizados pela Merrell Dow descobriram pequenas alterações nos membros de fetos de coelhas prenhes medicadas com Bendectin – alterações que os cientistas da companhia negligenciaram como insignificantes –, bem como abortos que o dr. Done crê que ocorreram porque os bebês sofreram malformações; que estudos *in vitro* realizados pelos *National Institutes of Health* descobriram que o Bendectin interferiu no desenvolvimento das células embrionárias dos membros; e que os dados extraídos de um estudo epidemiológico realizado para a Merrell Dow pelo dr. Bunde e pelo dr. Bowles, quando ajustado no sentido de excluir sujeitos canadenses (que poderiam ter comprado o medicamento sem receita), revelou que mães que tomaram o Bendectin correm um risco 30% maior de ter um bebê malformado³. O dr. Done explicou que sua convicção de que os defeitos de nascença de Mary Oxendine foram causados pelo Bendectin que sua

1 William Whewell, *Philosophy of the Inductive Sciences* (1840), in *SELECTED WRITINGS OF WILLIAM WHEWELL* 121-259, 257 (Yehuda Elkana ed., (1984)). (A palavra inglesa “consilience”, que creio que foi introduzida por Whewell, deriva do latim “con” e “silire”, “saltar junto”).

2 Oxendine v. Merrell Dow Pharm., Inc., 506 A.2d 1100 (D. C. 1986), devolvido, 563 A.2d 330 (D. C. 1989), cert. denied 493 U.S. 1074 (1990), devolvido, 593 A.2d 1023 (D. C. 1991) (“Oxendine 1991”), devolvido, 649 A.2d 825 (D. C. Ct. App. 1994) (“Oxendine 1994”), devolvido, Civ. No. 82-1245, 1996 WL 680992 (D. C. Super. Oct. 24, 1996). A descrição dos defeitos de nascença da sra. Oxendine provém de Oxendine, 506 A.2d p. 1103.

3 Oxendine, 506 A.2d p. 1104-09 (reportando parte do testemunho do dr. Done). O que apresentei no texto foi obviamente apenas um resumo muito esquemático do testemunho do dr. Done; ele ocupou o banco das testemunhas por três dias e meio e a transcrição do seu testemunho chega a quase 600 páginas. Id., p. 1108.

mãe tomou durante o período da gravidez em que os membros do feto estavam se formando estava baseada, não em algum desses estudos ou em qualquer uma dessas linhas de prova em si mesmas, mas em *todas* as várias peças de prova que ele atestou *tomadas em conjunto*⁴.

Em 1983, no primeiro julgamento, um júri adjudicou aos Oxendine 750.000,00 dólares de indenização por danos. Mas, revogando essa decisão e escrevendo que “está claro [...] que *não é possível tirar nenhuma conclusão em nenhum sentido de qualquer das bases de sustentação acima* com respeito a se Bendectin é um teratôgeno humano”⁵, a corte deferiu a moção do réu pleiteando julgamento sumário não obstante o veredicto. Os Oxendine apelaram e o Tribunal de Apelação reverteu a decisão e a devolveu com instruções para restabelecer o veredicto do júri, entendendo que o tribunal errou ao enfatizar o reconhecimento do dr. Done de que nenhum dos estudos por ele atestado seria individualmente suficiente para estabelecer a causa e “deixando de considerar [seu] testemunho de que todos os estudos combinados dão suporte a esse resultado”. O juiz adjunto Terry prosseguiu assim:

Como as peças de um mosaico, os estudos individuais mostram pouco ou nada quando vistos separadamente uns dos outros, mas eles se *combinaram para produzir um todo que era maior do que a soma de suas partes*: um fundamento para o parecer do dr. Done de que o Bendectin causou os defeitos de nascença do apelante⁶.

É claro que esse não foi o final da história do caso *Oxendine*: de fato, o caso chegou ao Tribunal de Apelação da Região de Colúmbia mais três vezes antes de finalmente ser resolvido em 1996. Quando o caso foi devolvido à primeira instância, a Merrell Dow apresentou moção para um novo julgamento, alegando que o dr. Done não havia apresentado apropriadamente as suas credenciais⁷; e em 1988 essa moção foi concedida. Os Oxendine apelaram

4 *Oxendine*, 506 A.2d p. 1108 (reiterando que o “[dr. Done] admitiu sua incapacidade de concluir que o Bendectin é um teratôgeno com base em qualquer dos estudos individuais que ele discutiu, mas também deixou claro que todos esses estudos devem ser vistos em seu conjunto e que, vistos dessa maneira, eles sustentam a sua conclusão”).

5 *Oxendine*, 506 A.2d p. 1103 (grifos meus).

6 *Oxendine*, 506 A.2d p. 1110 (grifos meus) (determinando que o julgamento sumário do tribunal de primeira instância laborava em erro porque, quando a prova era vista em seu conjunto, não era apropriado concluir que nenhum jurado sensato decidiria favoravelmente ao apelante).

7 *Oxendine*, 563 A.2d p. 332 (reportando que, em 3 e 11 de maio de 1983, o dr. Done testemunhou dizendo que era membro da *Wayne State Medical School Faculty*, quando na verdade ele apresentou carta de demissão no dia 24 de abril, a qual foi aceita pelo decano no dia 29 de abril; e listando quatro outros aspectos, em adição à sua posição na *Wayne State University*, com relação aos quais o dr. Done apresentou falsamente suas credenciais em juízo).

novamente e, em 1989, entendendo que o juiz de primeira instância errou ao conceder novo julgamento, o Tribunal de Apelação voltou a reverter a decisão, ordenando uma vez mais ao tribunal de primeira instância que restabelecesse o veredicto original⁸. De volta ao tribunal de primeira instância, os Oxendine entraram com um pedido para que o tribunal confirmasse o veredicto, mas a Merrell Dow apelou mais uma vez; e, em 1991, o Tribunal de Apelação sentenciou que o tribunal de primeira instância não poderia dar entrada a um julgamento final e inapelável referente a indenizações compensatórias enquanto não se concluísse o estágio das indenizações punitivas⁹. Em 1993, os Oxendine retiraram sua demanda por indenizações punitivas e entraram com moção para que fosse confirmado o veredicto referente às indenizações compensatórias; então a Merrell Dow requereu ao tribunal de primeira instância que o veredicto original fosse reconsiderado, dessa vez em razão de que novos estudos publicados desde o primeiro julgamento isentam o Bendectin. O tribunal de primeira instância, recusando-se a considerar esses novos estudos, uma sentença reafirmando o veredicto original do júri. A Merrell Dow apelou de novo; e, em 1994, reconhecendo que “reabrir a determinação de verdade científica do processo” estava em desacordo com a preocupação legal com a o caráter final do caso¹⁰, e, por conseguinte, estabelecendo um alto padrão de prova para que a Merrell Dow prevalecesse, o Tribunal de Apelação devolveu o processo uma vez mais – como disse o tribunal, com relutância e evidentemente esperando que o caso fosse resolvido rapidamente a favor dos Oxendine¹¹.

8 *Oxendine*, 563 A.2d p. 331 (julgando que o juiz de moções não abusou de seu poder discricionário ao julgar que a moção de suspensão foi apropriada, mas que ele errou ao suspender o julgamento e conceder novo processo). Id., p. 338 (revertendo a decisão e ordenando que a corte de primeira instância restabeleça o veredicto do júri).

9 *Oxendine*, 593 A.2d p. 1023 (revertendo concessão de indenizações compensatórias antes da conclusão do estágio de indenizações punitivas do processo).

10 *Oxendine*, 649 A.2d p. 831-32 (enfatizando a importância da finalidade no sistema legal). Ver também Susan Haack, “*Diferenças irreconciliáveis? O conturbado casamento entre ciência e direito*”, neste volume, pp. 173-201 (argumentando em parte que há uma tensão entre o falibilismo de final aberto da investigação científica e o desiderato legal da finalidade).

11 *Oxendine*, 649 A.2d p. 827 (esclarecendo que “somos relutantemente compelidos a devolver o processo para consideração limitada adicional”); ver também id., p. 834 (Schwelb, J., concordando) (o juiz Schwelb comenta que “as dilações até agora [...] já provocaram dano intolerável. [...] Não estamos em 1982 ou 1984 nem mesmo em 1990. [...] Tendo em vista o ponto a que chegamos hoje, as considerações de finalidade se tornaram tão urgentes que [...] nada menos que um apcrt extraordinariamente persuasivo da parte da Merrell Dow poderia justificar [...] postergação adicional ao ressarcimento da sra. Oxendine”).

Porém, em 1996 – passando a levar em conta os novos estudos apresentados pela Merrell Dow¹², as decisões em numerosos outros casos referentes ao Bendectin já conclusos desde o julgamento original¹³ e ações da FDA¹⁴ e do governo canadense¹⁵ – o tribunal de primeira instância julgou que o alto padrão exigido fora satisfeito e decidiu a favor da Merrell Dow¹⁶.

- 12 *Oxendine*, 1996 WL 680992, p. *14-21 (reportando que a Merrell apresentou 2 meta-análises de dados epidemiológicos referentes ao Bendectin posteriores a 1983 (Einarson et al., 1988; McKeigue et al., 1994) e 14 estudos epidemiológicos (Golding, 1983; Mitchell, 1983; Zieler, 1983; Asleton-Jick, 1984; Hearey, 1984; McCredie, 1984; Winship, 1984; Elbourne, 1985; Jedd, 1988; Aselton-Jick, 1985; Shiono, 1989; Erickson, 1991; McDonald, 1991; Khoury, 1994). Os advogados da demandante argumentaram que esses estudos, quando eram relevantes, eram falhos; por exemplo, que o estudo de Erickson do ano de 1991 omitiu salvaguardas cruciais como “períodos críticos” (presumivelmente o período da gestação em que as pessoas tomaram o Bendectin), mas o tribunal minimizou essas críticas como “críticas de advogado a um estudo científico, mais do que um estudo científico ou uma avaliação de especialista em contrário”. Id., p. *15 (citando e descartando os argumentos advocatícios).
- 13 *Oxendine*, 1996 WL 680992, p. *4-7 (listando oito casos federais conclusos a favor da Merrell Dow: *Lynch v. Merrell-National Labs.*, 830 F.2d 1190 (1st Cir. 1987); *Richardson v. Richardson-Merrell, Inc.*, 857 F.2d 823 (D.C. Cir. 1988), cert. denied, 493 U.S. 882 (1989); *Brock v. Merrell Dow Pharm., Inc.*, 874 F.2d 307 (3d Cir. 1989); *De Luca v. Merrell Dow Pharm., Inc.*, 911 F.2d 941 (3d Cir. 1990); *Ealy v. Richardson-Merrell, Inc.*, 899 F.2d 1159 (11th Cir. 1990), cert. denied, 498 U.S. 950 (1990); *Wilson v. Merrell Dow Pharm., Inc.*, 893 F.2d (10th Cir. 1990); *Turpin v. Merrell Dow Pharm., Inc.*, 959 F.2d 1349 (6th Cir. 1992), cert. denied 506 U.S. 826 (1992); *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311 (9th Cir. 1995). O tribunal também menciona *Blum e Havner*, mas observa que ambos se encontram em apelação. Ambos acabaram sendo decididos a favor dos réus. Ver *Blum ex rel. Blum v. Merrell Dow Pharm, Inc.*, 746 A.2d p. 4-6 (Pa. 2000) (confirmando a decisão do Tribunal Superior a favor da Merrell); *Merrell Dow v. Havner*, S.W.2d 706 (Tex. 1997) (revertendo a decisão do Tribunal de Apelação e julgando a favor da Merrell). Os dois nomes da companhia ré – Richardson-Merrell, Merrell-Dow – refletem mudanças na propriedade durante o período relevante. Ver JOSEPH SANDERS, *BENDECTIN ON TRIAL* 213-14 (1998) (descrevendo a história da companhia); ver também MICHAEL GREEN, *BENDECTIN AND BIRTH DEFECTS* (1996) (descrevendo a história do litígio em torno do Bendectin).
- 14 *Oxendine*, 1996 WL 680992, p. *23 (referindo-se a uma monografia sobre medicamentos anti-histamínicos vendidos sem receita, publicada pela FDA em 1994, que examinou o succinato de doxilamina e concluiu que ele era suficientemente seguro para ser incluído como um componente de tais anti-histaminas).
- 15 *Oxendine*, 1996 WL 680992, p. *23 (reportando que, em 1988, os consultores do *Special Advisory Committee on Reproductive Physiology to the Health Protection Branch* [Comitê Consultivo Especial sobre Fisiologia Reprodutiva da Agência de Proteção à Saúde] do governo canadense concluiu que o Bendectin não deveria ser retirado do mercado canadense e que a etiqueta de advertência deveria ser modificada à luz da falta de provas de uma associação com defeitos de nascença). Porém, ver também id., p. 23 n. 45 (registrando que o advogado dos demandantes chamou a atenção para o fato de que os membros do comitê canadense “eram ligados à Merrell – um fato de que o governo canadense não tinha consciência”).
- 16 Id., p. 34. Ao apresentar a narrativa intrincada dessa saga legal de longa duração baseei-me em parte na história recontada em Joseph Sanders, *Science, Law and the Expert Witness*, L. & CONTEMP. PROBS. (2008). O prof. Sanders fez uma conjectura muito plausível no sentido de que a Merrell Dow despendeu tanto tempo e dinheiro em sua defesa no caso *Oxendine* “visando manter um registro irrepreensível no litígio em torno do Bendectin. Mesmo um único veredicto final em favor do demandante poderia tornar bem mais difícil pleitear julgamento sumário em outros casos”. Sanders, nota 13 acima, p. 30.

O outro caso mais conhecido referente ao Bendectin foi, é claro, *Daubert*, no qual a Suprema Corte dos EUA proferiu sua sentença histórica em 1993¹⁷; e que chegou à decisão final um ano antes do caso *Oxendine*, quando o juiz Kozinski confirmou o julgamento sumário do tribunal de primeira instância a favor da Merrell Dow¹⁸. Os defeitos de nascença de Jason Daubert eram similares aos de Mary Oxendine¹⁹, e seus pais, como os dela, acreditavam que esses defeitos foram causados pelo Bendectin; porém, *Daubert* seguiu um caminho legal diferente do seguido pelo caso *Oxendine*. Em 1989, o Tribunal Regional acatou a moção da Merrell Dow requerendo julgamento sumário após excluir os testemunhos especializados trazidos pelos Daubert, em razão de que a prova científica é admissível unicamente se for “suficientemente estabelecida para ser aceita de modo geral no campo ao qual pertence”²⁰. Ele decidiu que, como nenhum dos numerosos estudos epidemiológicos publicados havia encontrado uma associação estatisticamente significativa entre Bendectin e defeitos de nascença, os pareceres dos especialistas de Daubert não eram aceitos de modo geral no campo ao qual pertenciam e, em consequência, não eram admissíveis. O Tribunal de Apelação da 9ª Circunscrição confirmou a decisão, citando especificamente *Frye*²¹. E por ter se baseado em *Frye* – algo quase sem

- 17 *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993) (determinando o padrão para admitir o testemunho científico especializado em processos federais).
- 18 *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 269 (5th Cir. 1995) (confirmando o julgamento sumário a favor da Merrell).
- 19 Natalie Angier, *High Court to Consider Rules on Use of Scientific Evidence*, N. Y. TIMES, Jan. 2, 1993, p. 1 (reportando que “Jason Daubert, de San Diego, nasceu há 19 anos com apenas dois dedos na mão direita e sem um dos ossos do seu antebraço direito”).
- 20 *Daubert v. Merrell Dow Pharm., Inc.*, 721 F. Supp. 570, 572 (S. D. Cal., 1989) (citando *United States v. Kilgus*, 571 F.2d 508,510 (9th Cir. 1988)) (citando *United States v. Brown*, 557 F.2d 541, 556 (6th Cir. 1977)) (citando *Frye v. United States*, 293 F. 1013 (D. C. Cir. 1923)).
- 21 *Daubert v. Merrell Dow Pharm., Inc.*, 951 F.2d 1128, 1129-30 (1991) (citando *Frye*, 293 F. p. 1014).

precedentes num caso civil²² – a Suprema Corte concedeu uma *certiorari*, para determinar se *Frye* havia ou não havia sido suplantado quando as *Federal Rules of Evidence (FRE)* foram adotadas em 1975.

Um memorial do *amicus [curiae]* de Kenneth Rothman e outros epidemiologistas levantou diversas questões epistemológicas importantes; esses *amici* argumentam que as análises dos tribunais inferiores em *Daubert* atribuíram peso demasiado a se os estudos são estatisticamente significantes, superestimaram a importância da publicação avaliada pelos pares²³ – e, o mais importante no que tange aos nossos atuais propósitos, que eles “excluíram de antemão o uso de inferências válidas que podem ser obtidas da combinação de muitos estudos, mesmo que nenhum desses estudos por si só justificasse tais inferências”²⁴. Porém, a decisão da Suprema Corte – de que *Frye* foi suplantado, mas que a *FRE 702* ainda exigia que os tribunais examinassem o testemunho especializado presta-do quanto à confiabilidade e à relevância – não aborda esse tema. No entanto, a decisão do ministro Blackmun prosseguiu enfatizando que, ao examinar a confiabilidade, os tribunais não deveriam olhar as conclusões dos especialistas,

22 Kenneth J. Cheseboro, *Galileo's Retort: Peter Huber's Junk Scholarship*, 42 AM. U. L. REV. 1637, 1695 (1993) (reportando que “não havia nenhum caso decidido pelos tribunais de apelação federais anterior a 1975 que tenha aplicado a Regra *Frye* em qualquer tipo de caso civil. Em 7 de abril de 1993, somente três decisões dessas haviam sido reportadas, duas das quais foram decididas em 1991”). Essas três decisões foram *Barrel of Fun v. State Farm Fire & Cas. Co., Inc.*, 739 F.2d 1028, 1031 (5th Cir. 1988); *Christopherson v. Allied-Signal Corp.*, 939 F.2d 1106, 1115-16 (5th Cir. 1991), *cert. denied* 112 S. Ct. 1280 (1991); *Daubert*, 951 F.2d 1106, 1115-16 (5th Cir. 1991). Pode-se questionar se *Christophersen* realmente “está baseado” em *Frye*, já que o tribunal lista quatro considerações, sendo *Frye* apenas uma delas. Ver *Christopherson*, 939 F.2d p. 1110. No entanto, quando a Suprema Corte negou a *certiorari* em 1992, o ministro White, juntamente com o ministro Blackmun, discordaram, argumentando que a pergunta se *Frye* foi suplantado pela *FRE* “é uma questão importante e recorrente”. *Christophersen v. Allied Signal Corp.*, 503 U.S. 912, 912 (1992) (White, J., discordando) (sustentando que a *certiorari* deveria ser concedida). *Barrel of Fun*, que se baseia mais inequivocamente em *Frye*, foi um caso referente a uma fraude de seguro contra incêndio em que a prova excluída envolvia uma “avaliação de estresse psicológico” do testemunho prestado, o qual o tribunal sustentou ser essencialmente similar à prova do detector de mentiras, que foi o tipo de prova em pauta em *Frye*. Ver *Barrel of Fun, Inc.*, 930 F.2d p. 1029 (considerando a prova em pauta como essencialmente similar à prova excluída em *Frye*).

23 Ver também *Brief for Petitioner [Memorial pelo peticionário]*, Darryl E. Chubin, Ph.D. et al., como *amici curiae*, *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993) (tratando, em primeira linha, da avaliação por pares). Ver também “Avaliação por pares e publicação: lições para advogados”, pp. 221-249 neste volume (fazendo distinção entre os conceitos amplo e estrito de avaliação por pares e analisando seu papel em *Daubert* e em litígios subsequentes).

24 *Brief for Petitioners [Memorial pelos peticionários]*, Professor Kenneth Rothman et al., como *amici curiae*, *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 570 p. *10 (1993) (grifos meus).

mas a sua “metodologia”²⁵. E assim, quando *General Electric Co. v. Joiner*²⁶ chegou à Suprema Corte em 1997, a disputa acerca da questão do peso conjugado da prova causal combinada foi formulada nos termos das “metodologias” rivais dos especialistas das partes.

Robert Joiner, que havia trabalhado durante muitos anos como eletricista para uma municipalidade na Geórgia, foi diagnosticado com câncer de pulmão do tipo microcelular, em 1991; ele tinha apenas 37 anos de idade. Acreditando que o seu câncer tinha sido fomentado por sua exposição aos bifenilos policlorados (BPCs) que contaminam o óleo usado para prevenir superaquecimento nos transformadores que desmontava e consertava por exigência do seu trabalho, ele moveu um processo contra o fabricante, a General Electric (GE). Os advogados do sr. Joiner trouxeram várias testemunhas especializadas que se propuseram a prestar testemunho acerca de uma série de estudos toxicológicos, *in vitro*, *in vivo* e epidemiológicos; eles argumentaram que, tomado em conjunto, esse acúmulo de provas seria suficiente para estabelecer um nexos causal. Eles explicaram que esses especialistas usaram “uma metodologia de peso de provas” – a mesma metodologia que a *Environmental Protection Agency [Agência de Proteção Ambiental]* usou para calcular o risco carcinogênico e a mesma metodologia que os próprios especialistas da GE usaram nesse mesmo caso. A GE, no entanto, negando essa imputação, argumentou que aquilo que os advogados de Joiner apresentaram como metodologia científica respeitável, na verdade, nada mais era que a “falácia do feixe” [*faggot fallacy*]: o erro de supor que, se uma pilha de provas fracas for suficientemente grande, é magicamente transmutada em prova forte²⁷.

O Tribunal Regional, enfocando um por um (alguns d)os estudos individuais, aos quais recorreram os especialistas de Joiner, excluiu o testemunho especializado de Joiner como inadmissível e decidiu a favor da moção da GE por julgamento sumário. Porém, o Tribunal de Apelação reverteu a decisão, sustentando que sempre que – como neste caso – a exclusão do testemunho especializado é determinante para o desfecho, a revisão judicial do processo deveria ser especialmente rigorosa; além disso, o Tribunal julgou a metodologia dos especialistas de Joiner cientificamente aceitável:

25 *Daubert*, 509 U.S. p. 592-93 (aplicar a *FRE 702* requer “uma avaliação preliminar sobre se a argumentação ou a metodologia subjacente ao testemunho é cientificamente válida e se a argumentação ou a metodologia podem propriamente ser aplicadas aos fatos em pauta”).

26 Gen. Elec. Co. v. Joiner, 522 U.S. 136 (1997).

27 Ver *Brief for Petitioners* p. *49, Gen. Elec. Co. v. Joiner, 522 U.S. 136 (1997).

Pareceres de qualquer tipo são derivados de peças individuais de prova que, tomadas a si mesmas, podem não ser conclusivas, mas *quando vistas em sua integralidade constituem os blocos que compõem uma conclusão perfeitamente razoável*, suficientemente confiável para ser submetida a um júri [...]”²⁸.

A Suprema Corte concedeu *certiorari* para determinar o padrão de revisão para tais decisões sobre a admissibilidade de provas. Decidindo unanimemente que o próprio padrão da revisão constituiu abuso do poder discricionário, o Tribunal que julgou *Joiner* fugiu do argumento de *Joiner* referente à “metodologia de peso de provas” com a observação ligeira de que metodologia e conclusões “não são inteiramente distintas uma da outra” e que o tribunal pode concluir com sensatez que há “simplesmente uma lacuna analítica demasiado grande entre os dados e o parecer prestado”. E, então, revisando rapidamente (alguns dos) testemunhos que os especialistas de *Joiner* teriam dado se tivessem sido admitidos, o Tribunal julgou, quase por unanimidade, que o Tribunal Regional não abusou do seu poder discricionário ao excluir os especialistas de *Joiner*²⁹.

Porém, sobre este último ponto houve um voto contrário, o do ministro Stevens, que argumentou que teria sido melhor ter devolvido o caso ao Tribunal de Apelação para reconsideração sob o padrão de revisão apropriado. Ele ressaltou que os especialistas de *Joiner* fizeram referência a numerosos estudos, sendo que apenas um consta nos autos e apenas seis chegaram a ser considerados pelo Tribunal Regional; além disso, prossegue ele, a visão majoritária sobre a questão da confiabilidade, que exigiu que ela minimizasse a distinção entre metodologia e conclusões, é “possivelmente não fiel a [...] *Daubert*”³⁰. (Deveras: afinal, a distinção entre metodologia e conclusões, que a maioria um tanto despreocupadamente pôs de lado em *Joiner*, constituiu o ponto principal e central em *Daubert*.)³¹ E, a exemplo do Tribunal de Apelação, o ministro Stevens acredita que há mérito no argumento epistemológico dos especialistas de *Joiner*:

28 *Joiner v. Gen. Elec. Co.*, 78 F.3d 524, 532 (11th Cir. 1996) (grifos meus). Os advogados da GE alegam que “quase as mesmas palavras foram citadas pelos cientistas e acadêmicos palavras que violam a metodologia da ciência”. Ver Brief for Petitioners p. *49, *Gen. Elec. Co. v. Joiner*, 522 U.S. 136 (1997) (citando PETR SKRABANEK e JAMES McCORMICK, *FOLLIES AND FALLACIES IN MEDICINE* 35 (Prometheus Books 1990) (caracterizando a “falácia do feixe”).

29 Ver *Joiner*, 522 U.S. p. 146-47 (sustentando que abuso do poder discricionário é o padrão aplicável e que o Tribunal Regional não abusou de seu poder discricionário ao excluir os especialistas de *Joiner*).

30 Id., p. 152 (Stevens, J., em seu voto parcialmente contrário).

31 *Daubert*, 509 U.S. p. 595 (“É claro que o foco deve estar unicamente nos princípios e na metodologia, não nas conclusões que eles geram”).

Não é intrinsecamente “não científico” para profissionais experientes *chegar a uma conclusão mediante a ponderação de toda a prova científica disponível* – este não é o tipo de “ciência de má qualidade” com que *Daubert* estava preocupado. Afinal, como resalta *Joiner*, a *Environmental Protection Agency* (EPA) usa a mesma metodologia para calcular riscos, se bem que use um limite um pouco diferente [...]”³².

Naturalmente a questão se e, em caso afirmativo, como uma compilação de peças de prova pode atingir conjuntamente o grau exigido de prova, quando cada uma por si não é capaz de fazê-lo, é uma questão que surge reiteradamente em processos referentes a reparações em virtude de substâncias tóxicas, embora não seja usual que isso ocorra tão explicitamente como em *Oxendine* e *Joiner*³³. O quebra-cabeças epistemológico se evidencia de modo particularmente vívido no primeiro caso descrito aqui, no testemunho do dr. Done em *Oxendine*: ele reconhece que só a prova da atividade toxicológica estrutural não é suficiente para provar onexo causal; nem só a prova de estudos *in vitro*, nem só a prova de estudos com animais, nem só as suas reanálises estatísticas. Mas, juntem-se todas elas, prossegue ele, e de alguma maneira – pronto! – elas equivalem à prova da causa.³⁴ Mas *como* exatamente? Ele não diz; nem o diz o juiz Terry, a despeito de sua bela metáfora do mosaico. E, pelo que conheço, essa questão ainda está por ser satisfatoriamente resolvida. O propósito da próxima seção é preencher essa “lacuna analítica”³⁵.

32 *Joiner*, 522 U.S. p. 153 (grifos meus).

33 Mas cf., por exemplo, *Castillo v. du Pont de Nemours*, 854 So. 2d 1264, 1272 (Fla. 2003) (reportando que [o dr. van Velzen] afirmou repetidamente que usou a testagem *in vitro* como uma fonte de dados em conjunção com outros dados confiáveis para chegar à conclusão. Ele testemunhou que *a consideração do conjunto de todos os dados é uma prática científica comumente aceita*) (grifos meus)). Para constar, aponto que a Flórida é um Estado que aplica a Regra *Frye* e que o padrão de revisão para sentenças de *Frye* é *de novo*.

34 Ver *Oxendine v. Merrell Dow Pharm. Inc.*, 506 A.2d 1100, 1108 (D. C. Cir. 1986) (reportando que “durante todo o seu testemunho, [o dr. Done] repetidamente deixou dito que o seu parecer se baseou não em um estudo isolado ou algum tipo isolado de prova, mas em quatro diferentes tipos de dados científicos vistos em combinação”).

35 Esses exemplos serão recorrentes em todo o ensaio; assim, talvez seja necessário dizer de uma vez que meu argumento *não* é que o *Bendectin* causa defeitos de nascença nem: que os BPC causam câncer microcelular de pulmão – nem, é claro, que eles não causam. Mesmo que eu tivesse toda a prova – que obviamente não tenho –, eu não seria competente para fazer tais julgamentos.

2. O argumento epistemológico

A primeira coisa a observar é que, enquanto até agora abordamos essa questão epistemológica a partir da perspectiva da prova causal em casos referentes a delitos com substâncias tóxicas, onde ela se sobressaiu de modo especial em termos legais, ela é, na realidade, bastante geral, emergindo virtualmente em cada área de investigação³⁶.

Pense-se, por exemplo, naquele meteorito descoberto na Antártida em 1984 e em relação ao qual se acreditou, com base nos gases que emitiu quando aquecido, que tivesse vindo de Marte há cerca de quatro bilhões de anos. Um químico de Stanford descobriu que o meteorito continha moléculas de hidrocarbonetos aromáticos policíclicos (HAPs), que são encontrados não só em escapamento de óleo diesel e ferrugem, mas também em matéria orgânica em decomposição; e outros cientistas descobriram que os cristais de carbonato no meteorito tinham a forma de cubos e lágrimas, a exemplo dos formados por bactérias na terra. Em 1997, o dr. David Mackay, do *Johnson Space Center [Centro Espacial Johnson]*, estava apto a dizer, em entrevista à *Newsweek*, que “*nós temos estas linhas de prova. Nenhuma delas por si mesma é definitiva, mas tomadas em conjunto, a explicação mais simples é a de que antigamente havia vida em Marte*”³⁷; e, na medida em que mais provas foram agregadas em mais de uma década de pesquisa continuada, essa conclusão se tornou mais solidamente assegurada³⁸. Tampouco se trata aqui de um exemplo isolado; pelo contrário, com respeito a virtualmente qualquer alegação científica assegurada de algu-

36 Talvez o Prof. Rothman e seus colegas tenham percebido isso. Ver Rothman, nota 24 acima, p. *10 (comentando que “essa observação a partir do senso comum não é nova nem controversa”).

37 Sharon Begley & Adam Rogers, *War of the Worlds: There Are no Little Green Men on Mars. But There Are Some Very Hostile Fellows on Earth Debating Whether There was Life on the Red Planet*, NEWSWEEK, Feb. 10, 2006, p. 56 (grifos meus).

38 Ver Thomas H. Maugh III, *Probe Enters Mars Orbit*, L.A. TIMES, Mar. 11, 2006, p. A12 (reportando que não se sabe se alguma vez houve água em Marte e que um segundo meteorito marciano também contém algo que podem ser fósseis de Marte). Ver também Michael Hanlon, *Is This Proof of Life on Mars? The Meteorite That May Finally Have Resolved the Great Mystery*, DAILY MAIL, Feb. 10, 2006, p. 40.

ma importância – que o DNA é o material genético³⁹, por exemplo, ou que as espécies evoluíram através de um processo de seleção natural⁴⁰ – a prova é uma malha complexa de elementos entrelaçados.

Nesse tocante, basear-se em muitas linhas de prova entrecruzadas nem mesmo é algo restrito às ciências naturais. Pense-se, por exemplo, em um historiador que baseia em prova arqueológica ou em prova documental (e talvez também na teoria científica subjacente a técnicas para datar restos ou para identificar o papel em que ou a tinta com que o documento foi escrito) ou numa combinação de registros escritos e no depoimento de testemunhas ainda vivas. De fato, esse ato de basear-se em toda uma malha de provas é ubíquo – é a regra, não a exceção. É lugar-comum na vida cotidiana: quando, por exemplo, depois de ler uma matéria surpreendente em um jornal eu compro um jornal diferente ou assisto ao noticiário televisivo para checar se outras fontes a confirmam⁴¹. E esse ato de basear-se em uma combinação de linhas de prova é também familiar a muitos contextos legais: quando, por exemplo, pedimos a um júri para chegar a uma conclusão baseado no testemunho de testemunhas oculares, no depoimento de um psicólogo a respeito das circunstâncias em que testemunhas oculares são mais ou menos confiáveis, na prova forense e no depoimento sobre o índice de erros cometidos por esse laboratório, e assim por diante.

Pelo fato de a questão epistemológica em pauta ser bastante geral, necessitamos de uma resposta geral. E também porque a justificação é claramente uma questão de grau (como dei por assentado ao dizer que a hipótese de que

39 Em 1944, quando publicou o relatório dos experimentos que são hoje reconhecidos como tendo estabelecido que o DNA e só a proteína constitui o material genético, Oswald Avery relutou em tirar a conclusão por escrito e ela não foi aceita de modo geral depois dos experimentos de Hershey e Chase, publicados em 1952. Ver Oswald T. Avery et al., *Studies of the Chemical Nature of the Substance Inducing Transformation in Pneumococcal Types*, 79 J. OF EXPERIMENTAL MED. 137 (1944); A. D. Hershey & Martha Chase, *Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage*, 36 J. GEN. PHYSIOLOGY 39 (1952). Em 1953, quando Watson e Crick publicaram seu ensaio sobre a estrutura do DNA, o papel do DNA na hereditariedade era entendido apenas de maneira muito imperfeita e, de acordo com Crick, somente nos anos de 1980 a conclusão ficou solidamente estabelecida. Ver James D. Watson & Francis Crick, *Molecular Structure of Nucleic Acids: A Structure for Deoxyribonucleic Acid*, 171 NATURE 737 (1953); Ver também FRANCIS CRICK, *WHAT MAD PURSUIT: A PERSONAL VIEW OF SCIENTIFIC DISCOVERY* 7 (1988).

40 Um resumo útil dessa prova extraordinariamente extensa e variada pode ser consultado em *Understanding Evolution: Your One-Stop Source for Information about Evolution*, http://evolution.berkeley.edu/evolibrary/articles/0_0_lines_01 (último acesso: 8 ago. 2008), clicando em “What Is the Evidence for Evolution? [Qual é a prova para a evolução?]”

41 Numa recente visita à Espanha, intrigada com os seus nomes, comprei exemplares dos dois jornais publicados em Múrcia: *La Verdad* (“A verdade”) e *La Opinión* [“A opinião”]. Amigos me disseram que o *La Verdad* seria uma publicação muito conservadora, e o *La Opinión* de cunho mais liberal. Ambas traziam a mesma matéria na primeira página, a saber, a de uma mulher estrangulada no centro da cidade.

antigamente havia vida bacteriana em Marte era pouco assegurada há uma década e está significativamente mais bem assegurada agora), a resposta deve explicar, em primeiro lugar, que fatores determinam se e, em caso afirmativo, em que grau a prova justifica uma conclusão; e, em segundo lugar, sob que condições esses fatores operam de tal modo que aumentam o grau de justificação quando diversas peças de prova são combinadas. Minha resposta recorrerá à explicação da estrutura de prova e dos determinantes do grau de justificação que apresentei no artigo *Evidence and Inquiry*⁴² e que ampliei e aprimorei no artigo *Defending Science*⁴³.

A prova se ramifica mais ao estilo de entrada em um jogo de palavras cruzadas; e minha explicação é inspirada nessa analogia. O grau de razoabilidade de um item de palavras cruzadas depende do quanto ele combina com a pista dada e com qualquer uma dos itens já completos com os quais tem interseção; o grau de razoabilidade desses outros itens é independente do item em questão e do quanto já foi completado no jogo de palavras cruzadas. De modo similar, proponho que o grau de justificação de uma conclusão (ou, como podemos formular em termos mais idiomáticos, qual é o grau de probabilidade de que a prova torne a conclusão verdadeira) dependa de três fatores:

- (i) de quão forte é a conexão entre a prova e a conclusão – *sustentação oferecida*;
- (ii) de quão sólida é a prova mesma,⁴⁴ independentemente da conclusão – *solidez independente*;

e:

- (iii) de quanto da prova relevante está contido na prova – *abrangência*.

Faço constar que, embora falemos frequentemente de grau de sustentação oferecida em termos de o quanto esta ou aquela prova torna tal conclusão provável ou falemos do grau de justificação em termos de o quanto é provável que a conclusão seja verdadeira, trata-se de probabilidades *epistêmicas*; elas não podem propriamente ser interpretadas como probabilidades estatísticas. De fato, dado o caráter multidimensional dos determinantes da qualidade

42 SUSAN HAACK, EVIDENCE AND INQUIRY: TOWARDS RECONSTRUCTION IN EPISTEMOLOGY 73-94 (1993) (2nd ed. 2009).

43 SUSAN HAACK, DEFENDING SCIENCE – WITHIN REASON: BETWEEN SCIENTISM AND CYNICISM 57-91 (2003).

44 De fato, “sólido” nesse caso significa “justificado”; mas isso não leva a um círculo vicioso; no final, chegamos à prova experiencial, que não tem necessidade de justificação. Ver id., p. 67-68 para uma explicação mais completa.

probatória, não há garantia nem mesmo de um ordenamento linear dos graus de justificação, quanto menos de uma possibilidade realista de especificar números (significativos) deles⁴⁵.

Também faço constar que esses três fatores não são bem simétricos. A sustentação oferecida está diretamente correlacionada com o grau de justificação; isto é, quanto maior a sustentação oferecida pela prova de uma conclusão, tanto mais justificada está a conclusão (como um item de palavras cruzadas é mais razoável quanto melhor ele combina com a pista e os outros itens já completos). Porém, a conexão entre solidez independente e justificação é um pouco mais complicada. Quanto mais independentemente segura for a prova *favorável* à conclusão, tanto mais justificada a conclusão; mas quanto mais independentemente segura for a prova *contrária* à conclusão, tanto menos justificada a conclusão (como num jogo de palavras cruzadas, o fato de que nossa resposta a 4 vertical combina com nossa resposta a 2 horizontal é tanto mais assegurador quanto mais confiantes estivermos de que 2 horizontal está correto; mas se nossa resposta a 4 vertical *não* combina com 2 horizontal, isso é tanto menos preocupante quanto *menos* confiantes estivermos de que 2 horizontal está correto). De modo similar, quanto mais abrangente for a prova *favorável* à conclusão, tanto mais justificada a conclusão; porém, se tornar a prova mais abrangente também a torna menos positiva, o aumento de abrangência diminui o grau de justificação da conclusão (como completar mais itens do jogo de palavras cruzadas nos torna mais confiantes da correção dos itens preenchidos se todos combinam uns com os outros, mas mina a nossa confiança se isso introduz inconsistências).

Assim, a combinação de peças de prova justificará a conclusão em um grau mais elevado do que qualquer dos seus componentes isoladamente poderia fazer-se, mas somente se, combinar os vários elementos incrementar a sustentação oferecida, incrementar a solidez independente da prova favorável à conclusão e/ou incrementar a abrangência por meio da introdução de elementos adicionais não menos sustentadores.

*

45 Ver id., p. 75-76 para um argumento mais completo de por que as probabilidades epistêmicas não satisfazem os axiomas do cálculo matemático das probabilidades. A tese não é nova. Ver JOHN MAYNARD KEYNES, A TREATISE ON PROBABILITY 28 (1921) (argumentando que “nem mesmo está claro que somos sempre capazes de situar [as probabilidades epistêmicas] numa ordem de magnitude”); RICHARD VON MISES, PROBABILITY, STATISTICS AND TRUTH 97 (2ª rev. em inglês ed., 1928) (argumentando que “nossa teoria da probabilidade nada tem a ver com questões como ‘Há alguma probabilidade de a Alemanha ser envolvida numa guerra contra a Libéria?’”).

Se aplicarmos essa análise um tanto abstrata a um exemplo esquemático baseado nos tipos de acúmulo de provas tipicamente encontrados nos processos referentes a delitos com substâncias tóxicas e examinarmos o efeito de combinar provas em termos de sustentação oferecida, solidez independente e abrangência, veremos como combinar provas pode – como o ministro Stevens e o juiz Terry acreditaram que poderia – incrementar o grau de justificação da conclusão causal.

Supondo que a alegação em pauta é que exposição à substância S causa ou fomenta a doença D: por exemplo, que a exposição de uma mulher grávida ao Bendectin causa defeitos de nascença em seu bebê ou que a exposição aos BPCs fomenta o desenvolvimento de câncer do pulmão. A prova, P, relevante para a conclusão, C, pode incluir qualquer um dos ou todos os elementos e_1, e_2, \dots, e_n a seguir:

- prova-epidemiológica (a partir de testes clínicos ou estudos médicos) da incidência de D entre os que foram expostos a S, quando comparada com sua incidência entre os que não foram expostos a S;
- meta-análises de tais estudos epidemiológicos, que indiquem se um aumento, caso haja algum, do risco de D é sugerido por seus dados combinados;
- prova sobre se a incidência de D na população cai quando S é retirado do mercado (ou afastado de prédios ou o que seja);
- prova sobre quais são os componentes de S (digamos, a, b e c) e se a exposição a qualquer outra substância que contenha um ou mais desses componentes ou a produtos químicos do mesmo tipo genérico é associada ao aumento do risco de D;
- prova de estudos *in vivo*, que indiquem se animais deliberadamente expostos a S desenvolvem D ou precursores de D;
- prova de estudos *in vitro*, que indiquem se células ou embriões deliberadamente expostos a S desenvolvem D ou precursores de D;
- prova sobre se há qualquer mecanismo biológico mediante o qual a exposição a S (ou a a, b e/ou c) pode causar D ou se há razões para acreditar que S (ou a, b ou c) não poderia causar D.

Porém, a prova com respeito à conclusão causal pode incluir também uma boa porção de outros tipos de informação, tendo relação um pouco menos direta com ela:

- metaprova com respeito a todos os tipos de prova listados acima: por exemplo, prova sobre o que se requer de um estudo epidemiológico, toxicológico, *in vitro* ou *in vivo* bem projetado e bem executado (por

exemplo, que variáveis precisam ser controladas, etc.), e o que constitui uma meta-análise bem projetada e bem executada (por exemplo, o que determina quais estudos são suficientemente bons para serem incluídos numa meta-análise e quais é melhor desconsiderar);

- informações de fundo sobre que outros fatores (como, por exemplo, suscetibilidades genéticas) podem contribuir para o desenvolvimento de D;
- informações de fundo (ou conjectura) sobre que proporção de casos de D derivam de que espécies de causas conhecidas (ou suspeitas);
- teoria relevante nas áreas química, biológica, fisiológica, genética, etc. potencialmente relacionada com S ou com D⁴⁶;
- ideias sobre o que, em meio ao que ainda não é conhecido, pode ser considerado de modo razoável como potencialmente relevante para a etiologia de D.

E adicionalmente pode haver prova (metametaprova?) sobre as fontes de todos esses tipos de prova⁴⁷, indiretamente relacionada com sua credibilidade e, a partir daí, indiretamente, com a credibilidade de C:

- prova de que estudos relevantes foram publicados após serem submetidos a avaliação pelos pares em revistas renomadas, ou foram publicados por privilégio editorial em revistas de condição mais baixa, ou nem chegaram a ser publicados;
- prova acerca de quem realizou a pesquisa relevante: talvez o fabricante ou cientistas patrocinados pelo fabricante (e se a pesquisa foi paga a partir do orçamento que o fabricante mantém para a pesquisa ou da verba reservada para processos judiciais), ou cientistas universitários que recebem alguma bonificação do fabricante, ou cientistas independentes que não têm ligação nenhuma com qualquer das partes de um processo;

46 Por exemplo, num período tão tardio quanto o início dos anos de 1950, acreditava-se amplamente que nada prejudicial poderia atravessar a placenta da mãe e entrar no feto. A partir de 1955, no entanto, tornou-se conhecido que substâncias com um peso molecular menor do que 1.000 *poderiam* atravessar a placenta e ingressar no sangue do feto. ROCK BRYNNER & TRENT STEPHENS, DARK REMEDY: THE IMPACT OF THALIDOMIDE AND ITS REVIVAL AS A VITAL MEDICINE 12 (2001).

47 Pelo fato de operadores do direito não serem especialistas em epidemiologia, toxicologia etc., e não terem o tipo de conhecimento amplo de fundo requerido para fazer julgamentos judiciosos de plausibilidade, este tipo de prova (indireta, externa) provavelmente desempenha um papel mais significativo em contextos legais do que deveria, em termos ideais.

- sustentação prova de que essa testemunha é (ou não é) um depoente frequente em casos desse tipo, de que seu currículo mostra que ela é (ou não é) uma testemunha especializada profissional e não mais um cientista ativante; etc.
- prova (metametametaprova?) quanto a se estudos patrocinados por fabricantes tendem a ser mais favoráveis aos produtos destes do que estudos realizados de modo independente,⁴⁸ quanto a com que frequência ensaios submetidos a revisão paritária são retratados,⁴⁹ quanto a se ensaios em revistas menos renomadas são retratados com mais frequência do que os que foram publicados em fóruns mais prestigiosos, etc., etc.

P pode ser completa (isto é, incluir prova de todos os tipos listados) ou pode ser incompleta; e pode ser inteiramente positiva (isto é, em sustentação de C contra não-C), ou inteiramente negativa, ou mista. Por razões óbvias, nos casos que chegam ao tribunal, a prova usualmente é incompleta, mista ou, na grande maioria dos casos, ambas as coisas; porque, se ela fosse inteiramente inequívoca de um ou de outro modo, ou a demanda jamais teria sido apresentada, ou ela teria sido resolvida.

Nenhum elemento singular de um acúmulo de provas do tipo P será suficiente por si mesmo para estabelecer uma conclusão causal. Os efeitos de S em animais podem ser diferentes dos seus efeitos em humanos. Os efeitos de b quan-

48 De fato, muitos estudos de estudos confirmam que é significativamente mais provável que pesquisa sobre medicamentos ou aparelhos médicos patrocinada por companhias seja favorável aos produtos da companhia do que a pesquisa independente. Ver, por exemplo, Richard A. Davidson, *Sources of Funding and Outcomes of Clinical Trials*, 1 J. GEN. INTERNAL MED. 155 (1986); Paula Rochon et al., *A Study of Manufacturer-Supported Trials of Non-Steroidal Anti-Inflammatory Drugs in the Treatment of Arthritis*, 154 ANNALS INTERNAL MED. 157 (1994); Lee S. Friedman & Elihu D. Richter, *Relationship Between Conflict of Interest and Research Results*, 19 J. INTERNAL MED. 52 (2004). Enquanto os comentaristas legais tendem a se preocupar com a ciência induzida pelo litígio, não deveríamos esquecer que a ciência induzida pela publicidade também pode ser tendenciosa. Ver, por exemplo, Kevin P. Hill et al., *The ADVANTAGE Seeding Trial: A Review of Internal Documents*, 149 ANNALS INTERNAL MED. 251, 251 (2008) (argumentando que documentos internos mostram que o teste ADVANTAGE do Vioxx, feito pelo laboratório Merck, em 1999, foi "um teste seminal desenvolvido pela divisão de publicidade do laboratório Merck para promover a prescrição do Vioxx (rofecoxib) quando ele viesse a estar disponível [...] em 1999").

49 O serviço de indexação médica PubMed atribui um número, PMID (PubMed Identifier [Identificador da PubMed]), a cada artigo e com ele é possível pesquisar, por exemplo, por "retratação de publicação". Sobre a retratação de trabalhos fraudulentos, ver, por exemplo, Laura Bonito, *The Aftermath of Scientific Fraud*, 124 CELL 873 (2006); Harold C. Sox & Drummond Rennie, *Research Misconduct, Retraction, and Cleansing the Medical Literature: Lessons from the Poehlman Case*, 144 ANNALS OF INTERNAL MED. 609 (2006); Jennifer Couzin & Katherine Unger, *Cleaning Up the Paper Trail*, 38 SCIENCE 38 (2006).

do combinados com a e c podem ser diferentes dos seus efeitos isoladamente, ou quando combinados com x e/ou y.⁵⁰ Até mesmo um estudo epidemiológico mostre uma forte associação entre exposição a S e aumento do risco de D seria insuficiente por si mesmo: ele pode ter sido mal projetado e/ou mal executado, por exemplo (além disso, o que constitui propriamente um estudo bem projetado – por exemplo, que controles são necessários – depende de informação adicional sobre as espécies de fator que possam ser relevantes). E até mesmo um excelente estudo epidemiológico pode captar, não umnexo causal entre S e D, mas uma causa subjacente tanto à exposição a S quanto a D; ou possivelmente refletir o fato de que, nos estágios bem iniciais de D, as pessoas desenvolvem um desejo de S. Nem a prova de que a incidência de D diminuiu depois que S foi recolhido é suficiente por si mesma para estabelecer a causa – talvez a vigilância nos relatos de D tenha sido relaxada depois que S tenha sido recolhido ou talvez a exposição a X, Y, Z também foi reduzida, e um ou todos esses fatores somados causam D, etc.⁵¹

Porém, combinar provas, como no meu exemplo esquemático, pode ajudar a excluir explicações diferentes da explicação de que S causa D e assim justificar a conclusão mais solidamente. Para entender sob que condições P poderia justificar C em grau mais elevado do que tomando qualquer dos elementos e_1, e_2, \dots, e_n individualmente, necessitamos olhar para o efeito da combinação desses elementos sobre a sustentação geral oferecida por P, sobre a solidez independente de cada elemento de P e sobre a abrangência de P.

(i) *Sustentação oferecida*: o grau de sustentação oferecido por uma prova a uma conclusão depende, em palavras breves e aproximadas, de quão bem a proya e a conclusão se encaixam para formar um relato explicativo. A prova combinada desse modo sustentará uma conclusão melhor do que suas partes

50 Como aparentemente ocorreu com a talidomida, que foi descrita como composta de "dois compostos bastante inócuos". Ver Brynner & Stephens, nota 46 acima, p. 8 (citando o dr. Robert Brent, sobre o qual ver a nota 51 abaixo).

51 O dr. Robert Brent, o editor da revista *TERATOLOGY*, que repetidamente testemunhou a favor da Merrell Dow nos casos referentes ao Bendectin como especialista em "dados referentes a tendências seculares", enfatizou que, embora o Bendectin esteja fora do mercado estadunidense desde 1983, o índice de defeitos de nascença reportados permaneceu constante. Mas, com base em um argumento paralelo ao que está no texto, isso é insuficiente por si mesmo para excluir uma conclusão causal. E, de fato, sabemos que, após a retirada do Bendectin, alguns médicos começaram a prescrever a vitamina B6 e meio comprimido de Unisom contra enjoo matinal; e que o succinato de doxilamina, o ingrediente suspeito do Bendectin, também é um ingrediente do Unisom (e do Nyquil). Ver Janelle Yates, *Nausea and Vomiting of Pregnancy: Q&A with T. Murphy Goodwin*, 16 OBG MGMT. 54, 55 (2004), 55 (recomendando vitamina B6 e, se o vômito continuar, adicionando 12,5 mg de doxilamina, dividindo ao meio o comprimido de Unisom vendido sem receita médica). Sobre o Nyquil (bem como o Unisom), ver também Joseph Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases*, 46 STANFORD L. REV. 1, 10 (1993).

componentes individualmente se o relato explicativo resultante da conjunção de P e C for melhor do que o da conjunção de e_1 e C, se for melhor do que o da conjunção de e_2 e C, e assim por diante. O que torna o grau de sustentação propiciado a C por P maior do que o grau de sustentação propiciado a C por e_1 , maior do que o grau de sustentação propiciado a C por e_2 etc., é o *quão estreitamente seus componentes se imbricam* para compor um relato explicativo. Por exemplo, a prova de um mecanismo biológico pelo qual S pode causar D está imbricada com a prova epidemiológica do aumento do risco de D entre os que foram expostos a S para explicar um aspecto anteriormente não explicado da história; a prova de que S contém b e de que b está associado com o aumento do risco de D está imbricada com a prova epidemiológica de aumento do risco de D entre os que foram expostos a S para aprofundar uma explicação anteriormente superficial; e a teoria biológica, fisiológica, química etc. de fundo está imbricada com a prova dos riscos para humanos da exposição a S para ampliar o escopo de um relato explicativo anteriormente estreito.

Para que os elementos de P estejam mesmo imbricados, os mesmos termos (“S”, “b”, “D”, etc.) devem ocorrer do começo ao fim, como é o caso na minha lista esquemática; e os elementos estarão imbricados de modo tanto mais estreito quanto mais estritamente esses termos estiverem caracterizados, isto é, quanto mais específicos forem. Por exemplo, a sustentação combinada será mais reforçada se “D” for um “câncer microcelular do pulmão” do que se for um simples “câncer do pulmão” ou um “câncer”, ou será mais reforçada se for “defeitos de nascença de redução de membros” do que se for apenas “defeitos de nascença”; será mais reforçada se “b” for “sucinato de doxilamina” do que se for uma “anti-histamina” ou mais reforçada se for “Benlate”⁵² do que se for um “fungicida”; e assim por diante.

Os elementos de E também estarão tanto mais estreitamente imbricados quanto maior for a similaridade fisiológica entre os animais usados em algum estudo com animais e os seres humanos. Os resultados de testes com beija-flores ou rãs dificilmente se coadunariam de alguma maneira com a prova epidemiológica de risco para humanos, ao passo que os resultados de testes com camundongos, ratos, porquinhos-da-índia ou coelhos estariam mais estreitamente imbricados com tal prova, e os resultados de testes como primatas ainda mais estreitamente. É claro que “similar” deve ser entendido como elipse para “similar

52 Em *Castillo v. E.I. du Pont de Nemours*, Benlate foi o fungicida ao qual a sra. Castillo alegou ter sido exposta e que ela acreditou ter causado o defeito de nascença do seu bebê, a saber, olhos de proporções severamente diminuídas (microftalmia). *Castillo v. E.I. du Pont de Nemours & Co., Inc.*, 854 So. 2d 1264 (Fla. 2003).

nos aspectos relevantes”; e determinar *quais* são os aspectos relevantes pode depender, entre outras coisas, do modo de exposição: se os humanos são expostos a S por inalação, por exemplo, é importante que os animais de laboratório usados tenham tido um índice similar de respiração. (Às vezes, os próprios estudos com animais podem revelar diferenças relevantes; por exemplo, os ratos nos quais a talidomida foi testada eram imunes ao efeito sedativo que ela teve em humanos, o que deveria ter levantado a suspeita de que ratos tenham tido uma escolha insatisfatória de cobaia para essa droga.)⁵³ Reiterando, os resultados dos testes com animais estarão tanto mais estreitamente imbricados com a prova de risco para humanos quanto maior for a similaridade da dosagem de S implicada. (Um dos pontos fracos do testemunho especializado de Joiner foi que os estudos com animais que serviram de base implicaram a injeção de doses maciças de BPCs no peritônio de um bebê de camundongo, ao passo que o sr. Joiner ficou exposto a doses muito menores quando o óleo isolante contaminado respingou em sua pele e nos seus olhos.)⁵⁴ O momento da exposição também pode ser importante, por exemplo, quando a alegação em pauta é que a exposição de uma mulher grávida a S causa este ou aquele tipo específico de dano ao feto.

Uma vez mais, os elementos de P estarão tanto mais estreitamente imbricados quanto mais estritamente os estudos *in vitro* equivalerem às condições da exposição humana. Por exemplo, os demandantes em *Castillo v. du Pont* fizeram um grande esforço para mostrar que as células expostas ao Benlate nos estudos *in vitro* aos quais recorreram estavam numa condição tão similar quanto possível à exposição a que o bebê ainda não nascido da sra. Castillo alegadamente foi submetido quando sua mãe foi acidentalmente borrifada com o Benlate usado nos campos adjacentes⁵⁵.

53 Brynner & Stephens, nota 46 acima, p. 48. “Foi perturbador que humanos responderam à talidomida caindo em ‘sono profundo, natural’, mas os ratos não. [...] O fato de que não pôde ser encontrada nenhuma dose letal para ratos pareceu duplamente perturbador [...] os ratos simplesmente não estavam absorvendo o medicamento” (Richardson-Merrell era o distribuidor da talidomida nos EUA. Id., p. 39. O remédio era vendido como pílula para dormir. Id., p. 14. Mais tarde, porém, depois que o médico australiano dr. William McBride descobriu que ele ajudava a minorar o enjoo matinal, ele passou a ser prescrito para esse propósito. Id., p. 22. Mais tarde, o dr. McBride se tornou um herói por chamar a atenção para os perigos da talidomida; e, em seguida, ficou desacreditado depois que se descobriu que ele falsificou resultados de um estudo com animais no esforço de chamar a atenção para o que ele acreditava serem os efeitos teratogênicos do Bendectin. Id., p. 197-99, 27-29).

54 Gen. Elec. Co. v. Joiner, 522 U.S. 136, 144 (1997).

55 Ver, por exemplo, *Castillo*, 854 So. 2d p. 1274 (reportando que “o dr. Howard levou em consideração as roupas que Donna Castillo estava vestindo quando foi exposta, sua altura e peso para determinar a quantidade de pele exposta e usou os dados de DuPont para calcular a quantidade de benomil [o ingrediente suspeito no Benlate] que poderia ter sido absorvido e que teria passado pelo sistema dela”).

(ii) *Solidez independente*: combinar provas pode também reforçar a solidez independente (assim como o fato de que determinado item do jogo de palavras cruzadas se imbrica com outros que, por seu turno, se imbricam com outros e assim por diante, dá uma razão a mais para pensar que ele está correto). É certo que adicionar prova a partir de um estudo com animais não tornará um estudo epidemiológico falho menos falho, e adicionar a prova de um mecanismo fisiológico não converterá um estudo *in vitro* malfeito em um estudo mais rigoroso. (Esse parece ser o ponto que Skrabanek e McCormick tentam provar quando explicam que a “falácia do feixe” é falaciosa porque “um feixe de provas fracas permanece fraco”.)⁵⁶ No entanto, se adicionarmos a uma prova epidemiológica apenas fracamente assegurada de um aumento do risco de D entre os que foram expostos a S a prova adicional de que há um mecanismo biológico pelo qual S leva a D, justamente essa prova adicional reforça a solidez da conclusão do estudo epidemiológico. (De modo similar, se eu somar uma coluna de números e chegar a um resultado *n*, mas não tiver certeza de que minha resposta está correta por ter sido interrompido no meio do meu cálculo, mas pedir a alguém que verifique a aritmética e este encontrar o mesmo resultado, isso propriamente aumenta a minha confiança no resultado que obtive na primeira vez – mesmo que isso não altere o fato de eu ter sido interrompido).

(iii) *Abrangência*: é óbvio que E é mais abrangente do que qualquer dos seus componentes isoladamente; e isso pode reforçar o grau de justificação de C (assim como completar um novo item num jogo de palavras cruzadas de um modo compatível com os itens existentes dá à pessoa razão para ficar mais confiante quanto ao todo). Se, por exemplo, à prova epidemiológica que indica um aumento do risco de D entre aqueles que foram expostos a S (e_1) adicionarmos prova sobre a composição química de S e os efeitos fisiológicos nocivos dos seus componentes (e_2) e prova do mecanismo biológico pelo qual a exposição a S causa D (e_3), essa prova combinada justificará a conclusão causal num grau mais elevado do que qualquer dos componentes parciais dessa prova isoladamente. (A prova estatística de associação entre o ato de fumar e o câncer de

⁵⁶ Skrabanek & McCormick, nota 28 acima, p. 35; Ver também A. R. Feinstein, *Scientific Standards in Epidemiologic Studies of the Menace of Everyday Life*, 242 SCIENCE 1257 (1988).

pulmão⁵⁷ justificará uma conclusão causal num grau mais elevado se for combinada com a prova de um mecanismo causal; a prova de que as mulheres são mais suscetíveis do que os homens justificaria uma conclusão causal em grau mais elevado se for combinada com a prova do papel dos hormônios femininos na aceleração do processo.)⁵⁸ No entanto, o grau de justificação diminuirá em vez de aumentar se a prova adicional for negativa ou mesmo menos positiva do que o restante. Se, por exemplo, adicionarmos à prova oriunda de estudos com animais, que indicam um aumento do risco de D naqueles que foram expostos a S (e_1), a prova de que um estudo epidemiológico *não* descobre aumento do risco para humanos (e_2), o grau de justificação propiciado a C por essa prova combinada será mais baixo e não mais elevado.

*

O que acabei de oferecer foi uma análise teórica, uma caracterização abstrata dos determinantes da qualidade evidencial – uma análise forte o suficiente para,

- ⁵⁷ Cinco estudos publicados em 1950 são vistos agora como pioneiros: Richard Doll & Austin Bradford Hill, *Smoking and Carcinoma of the Lung: Preliminary Report*, 2 BRITISH MED. J. 739 (1950) (PMID 14772469); M. L. Levin et al., *Cancer and Tobacco Smoking: A Preliminary Report*, 143 JAMA 336 (1950) (PMID 15415261); C. A. Mill & M. M. Porter, *Tobacco Smoking Habits and Cancer of the Mouth and Respiratory System*, 10 CANCER RES. 539 (1950) (PMID 14772728); Schrek et al., *Tobacco Smoking as an Etiologic Factor in Disease. Part I: Cancer*, 10 CANCER RES. 49 (1950) (PMID 15398042); E. L. Wynder & E. A. Graham, *Tobacco Smoking as a Possible Etiologic Factor in Bronchiogenic Carcinoma: A Study of 684 Proved Cases*, 143 JAMA 329 (1950) (PMID 15415260). (Em 1953, já haviam sido publicados mais 13 estudos desse tipo.)
- ⁵⁸ Michaela Kreuzer et al., *Hormonal Factors and Risk of Lung Cancer in Women?*, 32 INT'L J. EPIDEMIOLOGY 263 (2003) (que sugere exatamente isso). Mas ver também Leno Thomas et al., *Lung Cancer in Women: Emerging Differences in Epidemiology, Biology, and Therapy*, 120 CHEST 370, 370 (2005) (“há evidências que sugerem haver diferenças na patogênese e possível aumento de suscetibilidade ao câncer de pulmão em mulheres”); International Early Lung Cancer Action Program Investigators, *Women's Susceptibility to Tobacco Carcinogens and Survival After Diagnosis of Lung Cancer*, 290 JAMA 180, 180 (2006) (“mulheres parecem ter maior suscetibilidade a carcinógenos do tabaco, mas possuem um índice mais baixo de desfechos fatais do câncer de pulmão em comparação com os homens”); Geoffrey C. Kabat et al., *Reproductive and Hormonal Factors and Risk of Lung Cancer in Women: A Prospective Cohort Study*, 120 INT. J. CANCER 2214, 2214 (2007) (“diversas linhas de prova sugerem que fatores endócrinos podem desempenhar um papel no desenvolvimento de câncer de pulmão em mulheres, mas a prova é limitada e inconsistente”); Diana C. Marquéz-Garbán et al., *Estrogen Receptor Signaling Pathways in Human Non-Small Cell Lung Cancer*, 72 STEROIDS 135, 136 (2007) (“o status do estrogênio parece ser um fator significativo no câncer de pulmão em mulheres [...]”); Patricia O'Keefe & Jyoti Patel, *Women and Lung Cancer*, 24 SEMINARS IN ONCOLOGY NURSING 3, 4 (2008) (“mulheres podem ser mais suscetíveis do que homens aos efeitos carcinogênicos dos carcinógenos pulmonares. [...] A pesquisa nessa área está em andamento e sendo intensamente debatida”); Neal D. Freedman et al., *Cigarette Smoking and Subsequent Risk of Lung Cancer in Men and Women: Analysis of a Prospective Cohort Study*, 9 LANCET 649 (2008) (que sugere que a alegação de que mulheres são mais suscetíveis do que homens é questionável).

como vimos, mostrar que a prova combinada de fato pode tornar uma conclusão causal mais justificada do que pode fazer qualquer dos seus componentes isoladamente. Todavia, ela não visa constituir um procedimento decisório para chegar a uma conclusão sobre a confiabilidade ou o contrário da prova causal (ou de outra). Não obstante, ela lança alguma luz sobre o auê em torno da “metodologia do peso da prova” em *Joiner*. Já deve ter ficado evidente que a acusação da GE de que os especialistas de *Joiner* incorreram numa falácia⁵⁹ ao suporem que a prova combinada proporcione maior justificação à sua conclusão causal do que seus elementos individuais se baseia num erro. Porém, igualmente deveria estar claro – embora isso talvez não seja tão óbvio – que o apelo de *Joiner* à “metodologia do peso da prova” é, em si mesmo, um pouco enganoso, ao menos se ele quer dar a entender que há algo como um algoritmo ou protocolo, algum procedimento mecânico, eficaz para calcular o valor combinado da prova.

Isso também é evidente quando se olham mais de perto as *Guidelines for Carcinogen Risk Assessment* [*Diretrizes para avaliação do risco carcinogênico*], da *Environmental Protection Agency* – EPA [*Agência de Proteção Ambiental*], do ano de 1986⁶⁰, às quais se referem os advogados de *Joiner*⁶¹. Essas diretrizes recomendam que “a questão da probabilidade de um agente ser um carcinógeno humano deveria ser respondida no quadro de um juízo em que se pese a prova”⁶²; no entanto, elas não usam a expressão “metodologia do peso da prova” nem oferecem qualquer coisa parecida com um algoritmo para determinar o peso combinado da prova. A seção intitulada “Categorização do peso total da prova para carcinogenicidade humana” simplesmente descreve como as categorias são determinadas: “(1) O peso da prova em estudos com humanos ou estudos com animais é resumido; (2) essas linhas de informação são combinadas para produzir uma atribuição experimental a uma categoria (ver Tabela 1); (3) toda informação relevante de apoio é avaliada para ver se a designação do peso total da prova precisa ser modificada”; o que equivale a pouco mais do que “examinamos toda a prova disponível e usamos nosso juízo para avaliar o que ela mostra”. A Tabela 1 – descrita apenas “para propósitos ilustrativos” – é um pouco mais específica: por exemplo, ela indica que uma substância é categorizada como um carcinógeno humano somente se houver prova epidemiológica “suficiente” e como um provável carcinógeno humano se houver prova epide-

59 Ver nota 30 acima com o texto associado.

60 Environmental Protection Agency, *Guidelines for Carcinogen Risk Assessment*, 51 Federal Register 3392, 34000 (Sept. 24, 1986) [doravante *EPA Guidelines*].

61 Gen. Elec. Co. v. *Joiner*, 522 U.S. 136 (1997) (Stevens, J., discordando) (cita *Brief for Defendants*, 4041).

62 EPA Guidelines, nota 60 acima, p. 33996.

miológica “limitada”, mas prova “suficiente” a partir de estudos com animais.⁶³ Porém, isso equivale a pouco mais do que exigir prova epidemiológica antes de incluir a substância na categoria de risco mais elevado – desde que essa prova epidemiológica seja “suficiente”.

As *Diretrizes* mais recentes da EPA, de 2005, incluem uma seção com o título curioso, mas revelador “Narrativa do peso da prova” [*Weight of Evidence Narrative*], que explica que a EPA ainda “ênfatiza a importância de pesar toda a prova para chegar a conclusões sobre o potencial carcinogênico de agentes para o ser humano”, mas, afasta-se da “abordagem passo a passo” das diretrizes de 1986, dá agora “um único passo integrativo”. Dados de estudos epidemiológicos geralmente têm a preferência, “mas todo o conjunto de informações [epidemiológica, *in vivo*, *in vitro*, toxicológica, etc.] [...] poderia proporcionar percepções valiosas”⁶⁴. Até esse ponto talvez não sejam muito mais úteis do que as diretrizes de 1986; mas à medida que se prossegue na leitura, há diversas observações que vale a pena fazer. Em primeiro lugar, essas diretrizes usam a mesma metáfora do “encaixar” que usei aqui de modo totalmente independente:

a narrativa explica os tipos de prova disponível e como eles se encaixam para tirar conclusões, e [...] aponta questões/pontos fortes/limitações significativas dos dados e das conclusões⁶⁵.

Em segundo lugar, elas dão como certo – exatamente como fiz aqui ao articular *em que grau* a prova justifica a conclusão, e se um acúmulo de provas justifica uma conclusão *em um grau mais elevado* do que seus componentes – que a justificação é uma questão de grau:

descritores [“carcinógeno humano”, “provável carcinógeno humano” etc.] representam pontos ao longo de um *continuum de prova*; [...] há gradações e casos limítrofes [...]⁶⁶.

E, em terceiro lugar, elas reconhecem a distinção ressaltada por mim entre probabilidades de frequência (como “a probabilidade de que um sueco selecionado aleatoriamente seja protestante é de *n*%” ou “a probabilidade de que um homem norte-americano de 60 anos de idade viva até os 75 é

63 Id., p. 34000.

64 Environmental Protection Agency, *Guidelines for Carcinogen Risk Assessment*, EPA/630P-03/001F, p. 1-11 (março de 2005).

65 Id., p. 1-12.

66 Id., p. 2-51 (grifos meus).

de $m\%$) e probabilidades epistêmicas (como em “é extremamente provável que BPCs sejam carcinogênicos”):

embora o termo “provável” possa ter uma conotação probabilística em outros contextos, seu uso como um descritor do peso da prova *não corresponde a uma probabilidade quantificável* referente a se o produto químico é carcinogênico⁶⁷.

Porém, quando chegam à questão central “o que determina o peso da prova?”, essas diretrizes retrocedem aos assim chamados “critérios de Bradford Hill”, extraídos da palestra agora clássica *The Environment and Disease [O meio ambiente e a doença]*, proferida por Sir Austin Bradford Hill em 1965⁶⁸. No entanto, não se trata de critérios para determinar a qualidade da prova em termos gerais, mas, como dá a entender o título dado por Hill, estão focados especificamente na prova causal médica (especialmente na prova relativa à exposição ocupacional); além disso, eles se aplicam unicamente numa situação em que já há prova estatística de aumento do risco de D entre os que foram expostos a S. O que Bradford Hill oferece é uma lista de nove aspectos de uma associação conhecida entre S e D que deveriam ser levados em conta quando se chega a uma conclusão referente a se a conexão é causal ou não:

- (1) *Intensidade*: isto é, qual a dimensão do aumento do risco de D naqueles que foram expostos a S;
- (2) *Consistência*: isto é, se a associação entre S e D foi observada por diferentes pessoas, em diferentes lugares e tempos e sob diferentes circunstâncias;
- (3) *Especificidade*: isto é, se a associação é especificamente entre dada substância ou exposição ocupacional e dada doença;
- (4) *Temporalidade*: isto é, se a exposição a S precede D (em vez de, por exemplo, estar associada com os estágios iniciais de D);
- (5) *Gradiente biológico*: isto é, se a incidência de D aumenta à medida que aumente a exposição a S;

67 Id., p. 2-53 (grifos meus).

68 Austin Bradford Hill, *The Environment and Disease: Association or Causation?*, 58 PROC. ROYAL SOC'Y MED. 205 (1965). De acordo com o *Federal Reference Manual on Scientific Evidence [Manual de referência federal para prova científica]* 376 (Federal Judicial Center, 2. ed. 2000), Bradford Hill estava amplificando critérios propostos pelo Ministério da Saúde dos EUA para avaliar a relação entre o fumo e o câncer de pulmão. U.S. DEP'T OF HEALTH, EDUC., & WELFARE, *Smoking and Health Report of the Advisory Committee of the Surgeon General* (1964).

- (6) *Plausibilidade*⁶⁹: isto é, se a hipótese causal combina com o conhecimento biológico corrente;
- (7) *Coerência*: isto é, a interpretação dos dados em termos denexo causal não deveria conflitar (“seriamente”) com fatos conhecidos sobre a história e a biologia da doença;
- (8) *Experimento*: isto é, se a incidência de D diminui se uma ação preventiva é realizada para reduzir a exposição a S;
- (9) *Analogia*: isto é, se há alguma similaridade com outros casos conhecidos de nexos causais.

Vale apontar que Bradford Hill reconhece que “nenhum dos meus nove pontos de vista é capaz de aportar prova indiscutível a favor ou contra a hipótese de causa e efeito, e nenhum deles pode ser requerido como um *sine qua non*”⁷⁰.

Difícilmente causará surpresa que esses “critérios de Bradford Hill” tenham se mostrado tão duráveis⁷¹, pois contêm uma boa dose de bom senso. Porém, eles não constituem realmente “critérios”, ao menos no sentido em que esse termo às vezes é entendido, ou seja, não como um procedimento decisório ou mesmo uma lista de checagem que poderia ser seguida mecanicamente. (O termo legal “indícios” talvez seja menos enganoso.) A prova pode, por exemplo, satisfazer alguns desses critérios e outros não, ou pode satisfazer alguns em alto grau e outros num grau baixo; e Bradford Hill nada diz sobre como aquilatar o êxito de um desses indícios contra a falta de êxito de outro, nem sobre como comparar hipóteses que se coadunam bem com este indício e mal com o outro e a outra que se coaduna mal com este e bem com aquele. Isso tampouco chega a surpreender. Porque, na verdade – como propõe o meu estudo teórico e é indicado pela curiosa palavra “narrativa” usada pela EPA –, estimar o valor da prova complexa é, em certo sentido, inevitavelmente uma questão de juízo; isso quer dizer que alguém com experiência no campo po-

69 Enquanto eu escrevia isso, a imprensa reportou que um novo estudo, que revela que Vytorin não é mais eficaz do que um placebo com respeito a certo problema das válvulas cardíacas, descobriu também um aumento do risco de câncer naqueles que tomam o medicamento; mas que pesquisadores “declararam a última descoberta como ‘implausível’ e provavelmente resultado do acaso”. Presumo que a razão que alegaram tenha sido que não há mecanismo conhecido que pudesse plausivelmente ser suposto para explicar tal nexos. Ron Winslow & Shirley S. Wang, *More Vytorin Bad News Hits Merck*, *Schering*, WALL ST. J., July 22, 2008, p. B1, B2.

70 Hill, nota 68 acima, p. 299.

71 Eles aparecem, por exemplo, não só nas diretrizes de 2005 da EPA (ver nota 64 acima), mas também em SANDERS, nota 15 acima, p. 55-56, no *Federal Reference Manual on Scientific Evidence*, nota 66 acima, p. 375-76; e em *Modern Epidemiology* 24-28 (2. ed. Kenneth J. Rothman & Sander Greenland eds., 1998).

derá ver que a alegação causal (para usar a palavra de Bradford Hill) é “plausível”, porque ele colocou na balança toda uma malha de conhecimentos e conjeturas de fundo – uma malha de conhecimento de fundo que, no entanto, não articulou plenamente e talvez nem pudesse fazê-lo. De fato, é precisamente por ser a estimativa da prova complexa uma questão de juízo nesse sentido que até especialistas com boa qualificação e altamente competentes podem discordar de maneira sensata; porque, a menos e até que a prova seja acachapante numa direção ou na outra, diferenças sutis no complexo inarticulado do conhecimento e da conjetura de fundo posto na mesa pelo cientista individualmente podem produzir estimativas diferentes.

Ainda assim, pode ser útil mapear a lista de “critérios” ou “indícios” um tanto assistemáticos de Bradford Hill à base da estrutura mais articulada do estudo que propus. Com efeito, “consistência” equivale ao reconhecimento de que a prova combinada oriunda de diversas fontes, desde que todas apontem para a mesma direção, incrementa a justificação da conclusão de que o nexos é causal. A “coerência”, a “plausibilidade biológica” e a “analogia” de Bradford Hill parecem corresponder aos tipos de prova incluídos na minha lista sob “conhecimento de fundo” – o que quer que se conheça sobre um mecanismo potencial ou um mecanismo pelo qual S pode vir a causar D, qualquer teoria biológica, fisiológica etc. com a qual poderia se coadunar a conclusão causal, e assim por diante. A “especificidade” corresponde à conexão que estabelecem entre o quão estreitamente os elementos de E se encaixam e quão estritamente “S” e “D” são especificados; e “temporalidade” corresponde ao fato de que uma associação entre S e D encontrada até mesmo em um excelente estudo epidemiológico poderia ser o resultado de uma causa comum de exposição a S e de D ou do fato de que a própria presença de D leva à exposição a S. E “experimento” corresponde ao que descrevi como prova quanto a se a incidência de D se modifica quando a exposição a S é deliberadamente reduzida (“dados referentes a tendências seculares”, como diz o dr. Brent); e o “gradiente biológico” está refletido ao menos em parte pelas minhas observações sobre a extensão e a maneira da exposição a S.

Ao passo que a “intensidade” da associação entre S e D, isto é, a dimensão do aumento do risco, tem seu lugar em minha proposta como elemento relevante para eliminar a possibilidade de que um risco aparentemente elevado seja resultado do acaso, ela é o *primeiro* item na lista de Bradford Hill. Ele recorre ao exemplo da incidência de câncer escrotal em limpadores de chaminé – reportando que, entre estes, já nos anos de 1920 a taxa de mortalidade devido ao câncer escrotal era 200 vezes maior do que a taxa entre aqueles que não es-

tavam expostos ao alcatrão ou a óleos minerais.⁷² Vale apontar para o fato de que a conexão com “alcatrão ou óleos minerais” já está embutida no exemplo de Bradford Hill e que a possibilidade de que homens nos estágios iniciais do câncer escrotal tenham sido de algum modo atraídos para o exercício da profissão de limpador de chaminés parece tão remota que poderia ser negligenciada. Porém, a razão principal desse fator ser menos proeminente no meu exemplo esquemático do que na lista de Bradford Hill é simplesmente que os tipos de processos que chegam ao tribunal certamente não serão aqueles em que a associação é *tão* forte que a inferência de uma conclusão causal está virtualmente garantida, mas é muito mais provável que sejam aqueles em que, depois que o medicamento testado inclusive em testes clínicos amplos entrou no mercado, um número muito maior de pessoas o tomam e há evidência, ou suspeita, de que pode haver riscos não previstos para algum subgrupo da população.

3. Respondendo algumas questões controversas

O aparato teórico agora a postos propõe (ao menos incipientes) respostas a uma série de questões epistemológicas que frequentemente causaram confusão em litígios provocados por delitos com substâncias tóxicas – questões referentes à prova da causa geral (o tópico principal aqui) e até algumas questões referentes à prova da causa específica.

72 Hill, nota 68 acima, p. 295 (citando Richard Doll, *Cancer, in MEDICAL SURVEYS AND CLINICAL TRIALS: SOME METHODS AND APPLICATIONS OF GROUP RESEARCH IN MEDICINE* 333, 333 (2. ed. L. J. Wits ed., 1964). (De acordo com o dr. Doll, em 1775 Percivall Pott reportou que “câncer do escroto era caracteristicamente uma doença de limpadores de chaminés”; e, em 1933, J. W. Cook et al. provaram que “o 3,4-benzopireno foi responsável pela ação carcinogênica do piche na pele de animais”. Id., p. 333 (citando J. W. Cook et al, *The Isolation of a Cancer-Producing Hydrocarbon from Coal Tar*, 1 J. CHEM. SOC. 395 (1933)).

A prova epidemiológica de aumento do risco de D entre aqueles que foram expostos a S é essencial para provar a causa geral?⁷³ As diretrizes de 1986 da EPA observam que “estudos epidemiológicos proveem informação única sobre humanos que foram expostos a carcinógenos suspeitos”.⁷⁴ Elas prosseguem dizendo que “estudos epidemiológicos descritivos são úteis na geração de hipóteses e na provisão de dados de apoio”, mas “raramente podem ser usados para fazer uma inferência causal”; no entanto, estudos “analíticos” de controle de casos ou de populações “são especialmente úteis para avaliar riscos a humanos expostos”⁷⁵. Obviamente, estudos epidemiológicos bem projetados e bem executados que mostrem um aumento do risco elevariam significativamente o grau de justificação de uma conclusão causal; e é claro que, diferentemente dos estudos com animais, nos quais sempre se coloca a questão se os animais usados são suficientemente similares aos seres humanos nos aspectos relevantes, os estudos epidemiológicos envolvem sujeitos humanos (indubitavelmente a razão pela qual a Tabela 1 das diretrizes de 1986 da APA efetivamente permite que estudos epidemiológicos superem os estudos com animais). Não obstante, se houver prova positiva suficiente de outros tipos, uma conclusão causal pode ser justificada em um grau não negligenciável até mesmo na ausência de uma prova epidemiológica.

Isso é particularmente significativo quando, por uma razão ou outra, não houver estudos epidemiológicos relevantes disponíveis ou estes não forem

73 *Daubert v. Merrell Dow Pharm., Inc.*, 721 F. Supp. 570, 575 (S. D. Cal. 1989) (sustentando que, dado que havia um vasto corpo de prova epidemiológica referente ao Bendectin, parecer de especialista que não estivesse baseado em prova epidemiológica não era admissível. Ver, por exemplo, *Grimes v. Hoffman-LaRoche, Inc.*, 907 F. Supp. 33, 35 (D. N. H. 1995) (excluindo o testemunho do dr. Lerman de que Acutane desempenhou algum papel no desenvolvimento de catarata por parte do sr. Grimes em parte em razão de que, “em vez de basear-se em dados epidemiológicos, o dr. Lerman baseou seu parecer referente à causa geral, em primeira linha, em teoria científica, em experimento *in vitro* e naquilo que ele considera certos fatos científicos ‘aceitos de modo geral’”); *Sutera v. Perrier Group of Amer.*, 986 F. Supp. 655 (D. Mass. 1997) (excluindo o testemunho especializado dos demandantes porque eles “não apresentaram estudos epidemiológicos científicos submetidos à avaliação por pares que pudesse associar LPA [leucemia promielocítica aguda] [...] com exposição ao benzeno” em níveis relevantes); *In re. Rezulin Prods. Liab. Litig.*, 369 F. Supp. 2d 398, 411 (2005) (excluindo o testemunho especializado dos demandantes no sentido de que o medicamento para diabete Rezulin causou dano “silencioso” ao fígado, em parte em razão de que “não há testes clínicos nem estudos epidemiológicos observacionais que apoiem a posição dos demandantes”); *In re Bextra and Celebrex Mktg. Sales Practices & Prod. Liab. Litig.*, 524 F. Supp. 2d 1166, 1175 (2007) (excluindo o testemunho especializado dos demandantes no sentido de que Celebrex poderia provocar efeitos cardiovasculares numa dosagem de 200 mg diários em parte em razão de que “não há testes controlados aleatórios nem metanálises de tais testes ou metanálises de estudos observacionais que revelem uma associação entre Celebrex 200 mg/d e o risco de ataque cardíaco ou derrame”).

74 *EPA Guidelines*, nota 60 acima, p. 33995.

75 *Id.*

possíveis⁷⁶. Michael Gottesman argumenta que “é bastante raro” que “uma prova epidemiológica humana conclusiva esteja disponível”⁷⁷; porque quando se suspeita que um medicamento ou produto químico possa ser nocivo, é mais provável que os fabricantes instituíam “procedimentos de proteção para o uso futuro do produto” ou então o retirem do mercado, o que torna tal trabalho epidemiológico muito mais difícil de ser realizado. Por exemplo, continua ele, os BPCs vinham sendo usados rotineiramente em transformadores elétricos até que relatórios começaram a vinculá-los com certos tipos de câncer, até serem proibidos em 1977⁷⁸; depois disso, eles não foram mais usados em transformadores, não havendo mais nenhuma possibilidade realista de realizar estudos epidemiológicos sobre um possível vínculo entre BPCs e o tipo de câncer que o sr. Joiner desenvolveu⁷⁹.

De qualquer modo, é importante ter claro que “não há prova epidemiológica de aumento do risco de D naqueles que foram expostos a S” não é equivalente a “há prova epidemiológica de que não há aumento do risco de D entre aqueles que foram expostos a S”. (Diferentemente da assim chamada “falácia do feixe”, confundir essas duas proposições bem diferentes *seria* realmente uma falácia.) Por exemplo, no começo não havia prova nem para um lado nem para o outro de que pacientes que tomaram o Vioxx por menos de 18 meses experimentavam um aumento do risco cardiovascular – e no começo do litígio em torno do Vioxx, o laboratório Merck argumentou como se isso fosse uma prova de que não havia aumento do risco entre pacientes que tomaram o medicamento somente por um curto período⁸⁰. Porém, quando estudos posteriores

76 *Ver Castillo v. du Pont de Nemours*, 854 So. 2d 1264, 1269-70 (Fla. 2003) (reportando que o especialista do demandante argumentou que “estudos epidemiológicos clínicos não estão disponíveis porque Benlate é um produto químico tóxico e, por conseguinte, não apropriado para experimentos com humanos” e que “em casos em que uma exposição inicial é muito rara, há problemas inerentes com estudos epidemiológicos porque um cientista não pode [eticamente] expor um ser humano a um teratogêno conhecido a fim de estudar os efeitos”).

77 Michael H. Gottesman, *From Barefoot to Daubert to Joiner: Triple Play or Double Error?*, 48 ARIZ. L. REV. 753, 767 (1998) (O sr. Gottesman usou *Daubert* e *Joiner* como argumento a favor dos demandantes na Suprema Corte.)

78 90 Stat. 2003, 2025 (1976), 15 U.S.C. § 605(e) (1976).

79 Gottesman, nota 77 acima, p. 767.

80 “Numa confissão que poderia minar uma de suas defesas centrais nos processos relativos ao Vioxx, o laboratório Merck disse ontem que errou quando reportou, no início de 2005, que um teste estatístico crucial havia mostrado que Vioxx causa problemas cardíacos somente após 18 meses de uso”. Alex Berenson, *Merck Admits a Data Error on Vioxx*, N. Y. TIMES, May 3, 2006, p. C1. Em *Plunkett v. Merck & Co. (In re Vioxx Prods. Liab. Litig.)*, 410 F. Supp. 2d 565, 596-67 (E.D. La. 2005), os demandantes entraram com uma moção para excluir o testemunho do Merck de que Vioxx provoca efeitos protrombóticos somente se tomado por 18 meses ou mais; mas ela foi negada em razão de que ambas as partes se basearam no mesmo estudo (o estudo APPROVE), ainda que o tribunal devesse se preocupar somente com a metodologia e não com as conclusões tiradas.

examinaram o uso de Vioxx por períodos curtos, eles encontraram prova que dá a entender que o risco aumentava desde a primeira dose⁸¹. Isso nos ensina a seguinte lição: que a ausência de prova de que p é apenas isto – uma ausência de prova; ela não é prova de que não-p.

Se houver estudos epidemiológicos relevantes e eles não descobrirem aumento do risco de D entre aqueles que foram expostos a S, isso fatalmente leva sempre e de modo inevitável a uma alegação de causa geral? Não, nem sempre, nem necessariamente. Se houver bons estudos, aí sim; mas se eles forem significativamente falhos de modo a tornar provável que tenham subestimado o risco, seus resultados negativos não levam fatalmente a tal alegação. Em *Blum v. Merrell Dow*, por exemplo, o dr. Shapiro, especialista do réu, reconheceu sob interrogatório cruzado que seu estudo epidemiológico misturou numa mesma amostra mulheres que tomaram o Bendectin durante o período da gravidez em que ocorria a formação dos membros do feto e mulheres que tomaram o medicamento somente depois que os membros já estavam formados, e desse modo pode ter subestimado qualquer aumento do risco de defeitos de redução de membros⁸². Ou, para recorrer a um exemplo mais recente, sabemos como o estudo VIGOR, o primeiro teste clínico amplo com o Vioxx realizado pelo laboratório Merck, observou atentamente os efeitos gastrointestinais do Vioxx por um período mais longo do que observou os efeitos cardiovasculares; como resultado disso, não conseguiu detectar um aumento estatisticamente significativo do risco de ataque cardíaco e derrame⁸³.

81 Patricia McGertigan & David Henry, *Cardiovascular Risk and Inhibition of Cyclooxygenase: A Systematic Review of the Observational Studies of Selective and Non-Selective Inhibitors on Cyclooxygenase*, 296 JAMA 1633 (2006) (citando estudos que detectaram aumento do risco já no uso inicial de Vioxx); W. A. Ray et al., *Cyclo-oxygenase-2 Selective Non-Steroidal Anti-Inflammatory Drugs and Risk of Serious Coronary Heart Disease*, 360 LANCET 1071 (2002); D. H. Solomon et al., *Relationship Between Selective Cyclooxygenase-2 Inhibitors and Acute Myocardial Infarction in Older Adults*, 109 CIRCULATION 2068 (2004); Linda Lévesque et al., *Time Variations in the Risk of Myocardial Infarction Among Elderly Users of Cox-2 Inhibitors*, 174 CAN. MED. ASS'N J. 1563, 1563 (2006).

82 *Blum ex rel. Blum v. Merrell Dow Pharms., Inc.*, 1996 WL 1358523, 33 Phila. Co. Rptr. 193, 215-17 (Pa. Ct. Comm. Pleas 1996). Ver também “Qual é o problema com a ciência induzida pelo litígio?”. pp. 250-281 neste volume (examinando o testemunho especializado do réu em *Blum*).

83 Ver David Armstrong, *Bitter Pill: How the New England Journal of Medicine Missed Warning Signs on Vioxx – Medical Weekly Waited Years to Report Flaws in Article that Praised Pain Drug – Merck Seen as “Punching Bag”*, WALL ST. J., May 15, 2006, p. A1, A10 (examinando problemas com o relatório feito pelo laboratório Merck sobre o estudo VIGOR); Ver também Haack, acima n. 23, p. 288 (discutindo problemas no estudo VIGOR).

É aceitável desconsiderar ou simplesmente e em princípio excluir estudos epidemiológicos cujos resultados não sejam estatisticamente significativos⁸⁴. Não. É certo que quanto menos robusto estatisticamente for um estudo, tanto menos ele contribui para justificar a conclusão causal. Porém, o ponto crucial é que significância estatística é uma questão de grau e que o grau de exclusão convencionalmente aceito nada mais é que uma convenção – um ponto de exclusão adotado pela comunidade científica relevante e mantido em nível elevado para assegurar que o risco de falsos positivos seja minimizado⁸⁵. Bradford Hill estava certo ao escrever, há quase meio século, que a ênfase na significância estatística, que naquela época estava em rápido crescimento, queria dizer que “com demasiada frequência [...] captamos a sombra e deixamos escapar a substância” quando “deduzimos ‘nenhuma significância’ de ‘nenhuma significância estatística’”⁸⁶. Porém, a tendência que ele deplorou tornou-se hoje uma prática firmemente estabelecida⁸⁷. E infelizmente, como observam Rothman *et al.* em seu memorial do *amicus [curiae]* em *Daubert*, “um jurado a quem se diz que um corpo de dados não é ‘estatisticamente significativo’ é levado a crer que os dados são sem valor”; e, como eles continuam, a expressão “talismânica” “estatisticamente significativo” pode criar a impressão totalmente enganadora de que dados estatisticamente significativos sejam infalíveis⁸⁸.

*Dunn v. Sandoz Pharmaceutical Corporation*⁸⁹ é especialmente fascinante por causa da confusão que faz nessa questão. Em resumo: a sra. Dunn processou a firma Sandoz, o fabricante, alegando que o Parlodel, medicamento antilactação produzido por ela, teria provocado nela o derrame pós-parto que tivera; mas o tribunal excluiu o seu especialista para a causa geral, o dr. Kulig, em razão de que seu testemunho seria insuficientemente confiável para estar

84 Ver *In re Bextra & Celebrex Mktg. Sales Practices & Prod. Liab. Litig.*, 524 F. Supp. 2d 1166 (N. D. Cal. 2007) (excluindo o testemunho dos demandantes acerca do risco de eventos adversos naqueles que tomam 200 mg de Celebrex por dia em parte em razão de que os estudos epidemiológicos não encontraram associação estatisticamente significativa). Ver também *Daubert v. Merrell Dow Pharms., Inc.*, 727 F. Supp. 570 (S.D. Cal. 1989) (mencionando isso como parte da razão para excluir o testemunho especializado dos Daubert).

85 Além disso, há diferentes maneiras de calcular significância estatística, as quais às vezes levam a resultados diferentes, e a própria escolha da maneira às vezes é objeto de controvérsia. Keith J. Winstein, *Boston Scientific Stent Study Flawed*, WALL ST. J., Aug. 14, 2008, p. B1, B6 (reportando uma controvérsia desse tipo).

86 Hill, nota 68 acima, p. 299-300.

87 Winstein, nota 85 acima (apontando que “revistas médicas tipicamente não querem publicar” estudos cujos resultados não sejam estatisticamente significativos).

88 Parecer do Professor Kenneth Rothman et al., como *Amici Curiae* em apoio a petiçãoários, *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 570 (1993), p. *4.

89 *Dunn v. Sandoz Pharms. Corp.*, 275 F. Supp. 2d 672 (M. D. N. C. 2003).

à altura das exigências de *Daubert*⁹⁰. O dr. Kulig testemunhou: “Acredito que a causa existe porque apliquei os critérios de Bradford Hill”⁹¹; mas o tribunal concorda com Sandoz que dr. Kulig aplicou *mal* aqueles critérios ao não se dar conta de que eles contribuem apenas quando já há prova epidemiológica de uma associação entre a substância e o distúrbio. Até aqui tudo bem. Mas então o tribunal inseriu tacitamente uma frase adicional: teria sido preciso que o dr. Kulig tivesse “tido um estudo *estatisticamente significativo* como ponto de partida para a aplicação dos critérios de Bradford Hill”⁹². O tribunal até pode ter suspeitado corretamente que a aplicação dos critérios de Bradford Hill pelo dr. Kulig fosse amplamente decorativa⁹³, e com certeza procedeu corretamente ao ressaltar que esses critérios pressupõem alguma prova de uma associação como seu ponto de partida; mas ele evidentemente *não* tinha consciência da atitude cética de Bradford Hill em relação à insistência na significância estatística.

90 *Dunn*, 275 F. Supp. 2d p. 681; Ver também *Soldo v. Sandoz Pharms. Corp.*, 241 F. Supp. 2d 434 (W. D. Pa. 2003) e *Caraker v. Sandoz Pharms. Corp.*, 172 F. Supp. 2d 1046 (S. D. Ill. 2001). O dr. Kulig também se propôs a testemunhar para a mesma finalidade, mas foi excluído. A propósito, os advogados da sra. Caraker compararam a prova do especialista trazida por eles com *a montagem das peças de um quebra-cabeça* para estabelecer a causa. *Id.*, p. 1048. (Antes de adotar a analogia do jogo de palavras cruzadas, opere por algum tempo com a analogia de Michael Polanyi, que compara o trabalho da ciência com a montagem de um enorme quebra-cabeça. MICHAEL POLANYI, *THE REPUBLIC OF SCIENCE: ITS POLITICAL AND ECONOMIC THEORY* 51-52 (1962), reimpresso in *KNOWING AND BEING* 49 (Marjorie Grene ed., 1969).

91 *Dunn*, 275 F. Supp. 2d p. 677.

92 *Dunn*, 275 F. Supp. 2d p. 680.

93 De qualquer modo, como argumentei acima, os “critérios” de Bradford Hill podem ser, quando muito, *indícios* de umnexo causal. (O testemunho do dr. Kulig dá a entender uma certa ambivalência nesse ponto: “a comunidade toxicológica, meus pares, usa Bradford-Hill extensivamente. [...] [No meu testemunho,] eu dei um passo a mais e apliquei à análise um critério publicado, aceito de modo geral. [...] E, na minha opinião, o critério de Bradford-Hill constitui uma metodologia científica aceita de modo geral para análise de reações adversas a medicamentos”; no entanto, ele também reconhece que “você pode interpretar a prova de maneiras diferentes.”) *Dunn*. 275 F. Supp. 2d p. 677-78. Ao ler seu depoimento nesse caso, percebi que ele escreve como se Bradford Hill tivesse providenciado uma lista de checagem, a qual ele percorre, comentando, por exemplo, que “este critério [temporalidade] está claramente preenchido” etc., ao passo que Bradford Hill parece estar bem consciente tanto de que muitos dos seus critérios podem ser preenchidos em graus variados quanto de que é necessário usar o bom senso ao decidir a respeito do grau de probabilidade de que uma relação seja causal. (Também aponto, para constar, que “critéria” é uma forma plural e não singular, como o dr. Kulig parece pensar.) *Affidavit of Kenneth Kulig*, M. D., FAACT, FACMT p. 27-30. *Dunn v. Sandoz*, 275 F. Supp. 2d 672 (M. D. N. C. 2003)(No. 1:98 CV 00912), 2000 WL 34616176.

*É apropriado desconsiderar (ou, em princípio, excluir) a prova obtida de estudos com animais?*⁹⁴ É claro que não. Obviamente tais estudos podem contribuir para a justificação de uma conclusão causal – tanto mais quanto mais bem projetados e mais bem executados forem, usando animais, dosagens, modos de administração, períodos de administração, etc. apropriados. É claro e não menos óbvio que sempre haverá a possibilidade de que animais sejam afetados adversamente por S sem que os humanos o sejam, e vice-versa⁹⁵; e se testes em animais bem projetados e bem executados mostrarem um aumento do risco de D com a exposição a S, mas estudos epidemiológicos bem projetados e bem executados não mostrarem aumento do risco de D em humanos expostos a S, suspeitaríamos com razão que possa haver diferenças fisiológicas relevantes, das quais ainda não temos consciência.

A prova epidemiológica de, no mínimo, o dobro do risco é (epistemologicamente) essencial para estabelecer nexos causal específico? – isto é, para ir além da alegação geral de que exposição a S às vezes causa ou fomenta D e chegar à alegação específica de que foi a exposição desta pessoa a S que causou ou fomentou essa D do demandante é necessário mostrar que a exposição a S dobra o risco de D? Esta é a exigência, por exemplo, da lei de Nova Jersey, tendo sido imposta pelo juiz Kozinski quando ele reexaminou *Daubert* por devolução para

94 Ver, por exemplo, *Metabolife Int'l v. Wornick*, 72 F. Supp 1160, 1169 (S.D. Cal. 1999) (excluindo a prova científica da Metabolife, em parte em razão de que, por uma questão legal, estudos com animais são inadmissíveis, “devido às incertezas geradas pelo ato de extrapolar dos efeitos em camundongos e ratos para humanos”). Em 2001, o Tribunal de Apelação dos EUA para a Nona Região revogou essa exclusão. Ver *Metabolife Intn'l v. Wornick*, 264 F. Rptr. 3d 832 (2001). Ver também *In re Silicone Gel Breast Implants*, 318 F. Supp. 2d 879 (C. D. Cal. 2004) (excluindo a prova dos demandantes oriunda de estudos com animais em razão de que “extrapolações de estudos com animais para seres humanos geralmente não são consideradas confiáveis na ausência de uma explicação científica crível quanto a por que essa extrapolação é justificada”) (citando *Hall v. Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1410 (D. Or. 1996)). Em *Joiner*, o Tribunal Regional concordou com a GE que os estudos com animais, nos quais os especialistas de Joiner se baseavam, eram inadequados para estabelecer que a exposição de Joiner aos BPCs promoveram o seu câncer; na apelação, os advogados de Joiner (imprudentemente) argumentaram como se a questão fosse esta: os estudos com animais como tais podem, alguma vez, constituir um fundamento apropriado para o parecer de um especialista? Ver *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 144 (1997).

95 “Usualmente pode-se confiar no fato de que um composto que causa um efeito numa espécie de mamíferos causará o mesmo efeito em outra espécie. Este é um princípio básico da toxicologia [...]” *Federal Reference Manual on Scientific Testimony*, nota 68 acima, p. 410. No entanto, estudos com animais têm duas desvantagens: a dificuldade em extrapolar para humanos porque as “diferenças em termos de absorção, metabolismo e outros fatores podem resultar em variação nas respostas entre as espécies”; e porque “as dosagens elevadas costumeiramente usadas em estudos com animais” deixam abertas questões referentes à relação entre dosagem e resposta em humanos. *Id.*, p. 409.

complementação por parte da Suprema Corte⁹⁶. Porém, isso se baseia numa confusão. A ideia por trás de tal exigência presumivelmente é que, somente se a exposição a S no mínimo dobrar o risco de D, podemos inferir que as chances de que o/a demandante que foi exposto/a a S e desenvolveu D desenvolveu D *porque* ele ou ela foi exposto a S. Mas essa ideia está baseada numa confusão de probabilidades estatísticas com possibilidades epistêmicas; e, pensando bem, fica claro que a duplicação do risco estatístico não é nem necessário nem suficiente para provar uma causa específica.

A prova epidemiológica de risco dobrado *não é suficiente* para estabelecer causa específica: em primeiro lugar, se o estudo que mostra uma duplicação do risco for mal projetado e mal executado, teremos apenas *uma baixa probabilidade epistemológica de uma probabilidade estatística maior do que 50%*; e, em segundo lugar, porque até um estudo bem projetado e bem executado também pode mostrar que aqueles sujeitos que desenvolvem D quando expostos a S têm algumas características em comum – talvez se trate de pacientes mais velhos mais do que de pacientes mais novos, ou de mulheres mais do que de homens, ou dos sedentários mais do que dos ativos – e nosso demandante pode ser uma mulher anciã e sedentária. E a prova epidemiológica de uma duplicação do risco também *não é necessária* para estabelecer a causa específica: em primeiro lugar, porque estudos que não conseguem mostrar uma duplicação do risco podem ser falhos – por exemplo, não levando em conta o período da gravidez em que os sujeitos foram expostos a S, ou não levando em conta o fato de que foram incluídos sujeitos que podem ter sido expostos a S presente em medicamentos para resfriados ou insônia⁹⁷; e, em segundo lugar, porque até um bom estudo epidemiológico, que indique, em alto grau de probabili-

96 Reexaminando *Daubert* por devolução para complementação por parte da Suprema Corte, o juiz Kozinski argumentou que os especialistas dos *Daubert* teriam de ser excluídos sob os novos padrões, assim como foram sob a Regra *Frye*; julgando que, a não ser que alegasse demonstrar que o Bendectin no mínimo dobra o risco de defeitos de nascença, um especialista teria de ser excluído em razão da irrelevância. Ver *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1320-21 (9th Cir. 1995) (“A lei referente ao delito civil da Califórnia exige que os demandantes demonstrem não apenas que o Bendectin aumenta a possibilidade de dano, mas que seja mais provável do que improvável que ele tenha causado os danos *por eles sofridos*” (citando *Jones v. Ortho Pharm. Corp.*, 163 Cal. App. 3d 396, 209 Cal. Rptr. 456 (1985)). O tribunal prosseguiu: “em termos de prova estatística, isso significa que os demandantes devem estabelecer [...] que [o fato de suas mães terem tomado o Bendectin] mais que dobrou” o risco. Ver *Daubert*, 43 F.3d p. 1320; Ver também *id.*, p. 1321 (citando *DeLuca v. Merrell Dow Pharm., Inc.*, 911 F.2d 941, 941, 958 (1990), onde a exigência da lei de Nova Jersey, a saber, que os demandantes devem demonstrar que é mais provável do que improvável que o Bendectin tenha causado os defeitos de nascença de Amy DeLuca, é interpretada no sentido de que “o risco relativo de defeitos de redução de membros que emergir dos dados epidemiológicos [...] deve, no mínimo, exceder ‘2’”).

97 Ver nota 51 acima e o texto associado.

dade epistêmica, que há uma duplicação do risco pode também indicar que aqueles sujeitos que desenvolvem D têm alguma característica (como ter mais de 50 anos de idade ou ser sedentário ou estar sujeito a alergias ou o que seja) que está ausente no demandante⁹⁸.

Há um problema relacionado com outro argumento às vezes encontrado, a saber: já que (digamos) se crê com base em prova confiável que 10% dos casos de D são genéticos e 20% são causados por fatores ambientais, ao passo que as causas dos restantes 70% são desconhecidas, as chances são de que essa D do demandante não foi causada, como alegado, pela exposição a S. Porém, aqui a confusão entre probabilidades estatísticas e epistêmicas é encoberta por confusões de dois outros tipos: uma falsa suposição de que a causa de D deva ser *ou* genética *ou* ambiental (quando pode muito bem haver interação entre as duas); e: tratar “desconhecidas” como se se referisse a outro tipo de causa, como “genética” ou “ambiental” – quando na verdade se trata obviamente de uma expressão de *desconhecimento*. Se um demandante argumenta que foi a exposição a S que fez com que ele desenvolvesse D e o réu responder que isso é improvável, por sabermos que 70% dos casos de D têm origem em causas desconhecidas, a resposta do réu é defectiva – porque se a alegação do demandante for verdadeira, o que *pensamos* que sabemos a respeito de que proporção dos casos de D é causada por fatores conhecidos, e que proporção é causada por fatores desconhecidos, pode não ser, no fim das contas, conhecimento genuíno.

Quando Donald Rumsfeld fez aquela notória observação sobre “desconhecimentos desconhecidos”⁹⁹, sendo a questão em pauta obviamente o serviço de inteligência referente ao Iraque, talvez eu tenha sido a única pessoa no país que não sorriu ironicamente; de qualquer modo, pelo menos de uma perspectiva estritamente epistemológica, o secretário Rumsfeld levantou um

98 Numa nota de rodapé, o juiz Kozinski reconhece esse problema pelo menos em parte: “não há dúvida de que haverá resultados injustos com esse padrão. Se um medicamento aumenta a possibilidade de defeitos de nascença, mas não chegar a dobrá-lo, alguns demandantes cujos danos são atribuíveis ao medicamento serão incapazes de ganhar a causa”; mas descarta isso com o comentário de que “há uma injustiça inversa sob um regime que permite que ganhe a causa todo aquele que pode ter sido afetado pelo medicamento” e que isso é um assunto que deve ser resolvido pelos Estados. *Daubert*, 43 F.3d p. 1320 n.13. Ele também reconhece a possibilidade de que possamos ter prova de que o demandante pertence a uma subclasse da população cuja suscetibilidade é maior do que a usual, mas aponta que os demandantes em *Daubert* não trouxeram prova com essa finalidade. Ver *id.*, p. 1321 n. 16.

99 Donald H. Rumsfeld, U.S. Sec’y of Def., Dep’t of Def., News Briefing (Feb. 12, 2002), <http://www.defenselink.mil/transcripts/transcript.aspx?transcriptid=2636> (último acesso em 23 de novembro de 2008) (“Sempre acho interessantes relatórios que dizem que alguma coisa não aconteceu. Porque, como sabemos, há conhecimentos conhecidos, há coisas que sabemos que sabemos. Também sabemos que há desconhecimentos conhecidos, quer dizer, sabemos que há coisas que não sabemos. Mas também há desconhecimentos desconhecidos – aqueles que nem sabemos que não sabemos”).

ponto genuinamente importante: nós não só podemos não ter toda a prova que sabemos que seria relevante (os “desconhecimentos conhecidos” da linguagem rumsfeldiana), mas também pode haver prova que não temos e que nem ao menos percebemos que é relevante. Isto – o problema rumsfeldiano dos “desconhecimentos desconhecidos” – também é relevante para a próxima questão na minha lista.

*Podemos inferir do fato de que as causas de D ainda são desconhecidas e de que um demandante desenvolveu D depois de ter sido exposto a S que foi esta exposição que causou a D da sra. X ou do sr. Y?*¹⁰⁰ Não. Tal prova certamente nos daria uma razão para examinar a possibilidade de que S seja a ou uma causa de D. Mas a conversa descuidada da “inferência à melhor explicação” mascara o fato de que o que no presente parece ser a explicação mais plausível pode não ser isso realmente – de fato, pode não ser explicação nenhuma. Podemos não conhecer todas as potenciais causas de D, nem mesmo saber que outras possíveis explicações seria recomendável investigar.

4. O argumento legal

Sob *Daubert* os tribunais devem examinar o testemunho especializado prestado¹⁰¹ em termos de relevância e confiabilidade (“probatória”). Vale a pena deter-nos por um momento para ressaltar que a relevância, assim como a confiabilidade, é uma questão factual. Isto é, se (e em que grau) p é relevante para q não é uma questão de pura lógica, mas depende de fatos a respeito do mundo: se e somente se a astrologia for verdadeira, por exemplo, a posição dos planetas na hora do seu nascimento é relevante para o modo como as coisas acontecerão com você nesta semana. Mas o ponto focal neste ensaio é a questão da confiabilidade.

100 Ver, por exemplo, *Rosen v. Ciber-Geigy Corp.*, 78 F.3d 316, 318 (7th Cir. 1996) (sustentando que o Tribunal Regional não abusou de seu poder discricionário ao excluir o testemunho do dr. Fozzard no sentido de que o ataque cardíaco do sr. Rosen foi causado pelo fato de ter usado um emplastro de nicotina nos três dias anteriores à ocorrência: “quando um evento inusual segue imediatamente a outro evento inusual a pessoa comum infere uma relação causal. [...] Porém, especulações leigas sobre causalidade médica, mesmo que sejam plausíveis, são uma base perigosa para inferir causalidade [...]”).

101 Há quem prefira formular isso de uma maneira um pouco diferente: que *Daubert* impôs claramente essa exigência com respeito ao testemunho científico, mas foi somente quando a Suprema Corte esclareceu o escopo de *Daubert* em *Kumho Tire* que ficou claro que a exigência também se aplicaria ao testemunho especializado distinto do científico. *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137 (1999).

No meu modo de ver, a confiabilidade é uma questão de grau; a admissibilidade, em contraposição (normalmente) é categorial: ou se permite que uma testemunha deponha ou não, ou se permite que ela deponha a respeito desta ou daquela questão¹⁰² ou não se permite¹⁰³. Assim sendo, o tribunal, ao determinar se o testemunho é ou não é admissível, impõe uma dicotomia bem nítida, uma dicotomia do sim ou não, a um *continuum* de graus de confiabilidade.¹⁰⁴ A discrepância entre a natureza categorial da admissibilidade e o caráter gradacional da confiabilidade tornou-se ainda mais acentuada a partir de 2000, quando a *FRE 702* foi reformulada para se exigir que o testemunho especializado se baseie em dados “suficientes”, “confiavelmente” obtidos e “confiavelmente” aplicados aos fatos em questão¹⁰⁵. E as complexidades são constituídas pelo fato de que uma parte, confrontada com a contestação de *Daubert* ao seu testemunho especializado prestado, deve mostrar “pela preponderância da prova” que esse testemunho satisfaz o padrão legal de confiabilidade. O que as partes devem mostrar, pelo visto, é que “*seja mais provável que improvável* que esse testemunho satisfaça de modo *suficientemente provável* a cláusula de confiabilidade de *Daubert*”. Bem: trabalho há muitos anos em epistemologia, mas preciso dizer que para mim o significado disso é um mistério.

Mas o problema mais imediatamente relevante aos nossos atuais propósitos é que *Daubert* parece impor uma espécie de atomismo probatório¹⁰⁶ que se

102 Ver, por exemplo, *U.S. v. Llera-Plaza*, Nos. CR 98-362-10, 98-362-11, 98-362-12 (E. D. Pa. Jan. 7 2002). O juiz Pollack decidiu que embora o testemunho de peritos em impressões digitais seja admissível em certos assuntos, “não se permitirá às partes apresentar depoimento que expresse parecer de um testemunha especializada de que uma impressão latente em particular coincide ou não coincide com as impressões registradas de uma pessoa em particular e, desse modo, é ou não é a impressão digital daquela pessoa”. Id., p. 19.

103 Porém, ver Transcript of Bench Ruling p. 1484, *U.S. v. Brown*, 05 Cr. 938 (JBR) Southern District Reporters 1468, 1484 (S. D. N. Y., 2008) (argumentando que a admissibilidade sob *Daubert* não precisa ser interpretada como categorial e permitindo que peritos em balística testemunhem somente no sentido de que suas conclusões são mais prováveis do que improváveis; e observando que o tribunal “teve uma discussão há cerca de um ano com o Prof. Dan Capra [das Escolas de Direito de Colúmbia e Fordham]”, perguntando-lhe se “a *FRE 702* pretendia ser uma regra absoluta, no sentido de ou está dentro ou está fora” e ele disse: não, de jeito nenhum. [...] É claro que é apenas a sua opinião. [...] Porém, há muitas situações em que você pode achar que a metodologia e o testemunho são confiáveis em certo grau, não em grau muito elevado – não no grau proposto por uma testemunha e, nesse caso, aquela *Rule 702* diz para não jogar a coisa toda fora, mas confiná-la dentro de certos limites”). Ver também *U.S. v. Glynn*, No. 06 Cr. 580, 2008 WL 4293317, p. *1 (S. D. N. Y. Sept. 22, 2008) (fazendo referência à sentença do tribunal em *Brown*).

104 Ver Dale Nance, *Two Concepts of Reliability*, (APA) NEWSL. ON PHIL. & L., Fall 2003 p. 123.

105 *Federal Rule of Evidence 702*.

106 Às vezes também chamado de “corpuscularismo”. Ver Thomas McGarity, *Our Science is Sound Science and Their Science is Junk Science: Science-Based Strategies for Avoiding Accountability and Responsibility for Risk-Producing Products and Activities*, 52 KAN. L. REV. 892, 921 (2004).

choca com o caráter mais holístico da prova da causa em sua maioria. O problema é bem perceptível em *Joiner*, quando a Suprema Corte examinou um por um (alguns dos) estudos que os especialistas de *Joiner* teriam citado se tivessem sido admitidos e julgou que nenhum deles estaria à altura de *Daubert*. Mas a sentença de 1995 do juiz Kozinski referente ao processo devolvido para complementação revela que o problema deriva do próprio processo *Daubert*. Pelo fato da lei ter mudado desde que o tribunal de primeira instância concedeu o julgamento sumário a favor da Merrell Dow em 1989, o juiz Kozinski argumentou que seria justificado conceder aos demandantes a oportunidade de demonstrar que o testemunho especializado por eles prestado satisfaz ao novo padrão¹⁰⁷; ele prosseguiu dizendo que, no entanto, não faria sentido proceder daquela maneira se já estivesse claro que os referidos especialistas teriam de ser excluídos sob *Daubert*, do mesmo modo que foram sob *Frye*. E isso, de fato, continuou ele, já estava claro. Examinando um por um os testemunhos prestados pelos especialistas dos *Daubert*, o juiz Kozinski observa, em primeiro lugar, que todos eles, exceto um, propuseram-se a testemunhar somente que havia uma possibilidade de que o Bendectin cause defeitos de nascença, e nem mesmo alegam, quanto menos demonstram, que o fato de uma mãe tomar o medicamento mais do que dobra o risco, e, diante disso, teriam de ser excluídos com base na cláusula de relevância¹⁰⁸; e, em seguida, que o dr. Palmer, o único especialista que alegou mais do que isso, a saber, que o Bendectin causou os defeitos de nascença de Jason Daubert, simplesmente não tinha uma metodologia decente, e, diante disso, teria de ser excluído com base na cláusula de confiabilidade.

E essa estratégia atomística está implícita na sentença do tribunal de *Daubert*, de acordo com a qual cada item da prova especializada deve ser examinado quanto à (relevância e) confiabilidade. Para ser admissível, e_1 deve ser (relevante e) confiável, e_2 deve ser (relevante e) confiável, e_3 deve ser (relevante e) confiável, e assim por diante¹⁰⁹. Porém, se meu argumento epistemológico estiver correto, a combinação de e_1, e_2, \dots, e_n pode justificar melhor a conclusão causal do que

107 Ver *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1315 (9th Cir. 1995). De fato, a retórica de *Daubert* 1993 foi que o novo padrão era mais receptivo à admissão de testemunho especializado do que a antiga e "austera" Regra *Frye*. Ver *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 578, 589 (1993) ("que o padrão austero, ausente do e incompatível com as *Federal Rules of Evidence* não deveria ser aplicado a processos federais").

108 Ver também nota 96 acima e o texto associado.

109 Ver McGarity, nota 106 acima, p. 924 ("sob a abordagem corpuscular, ou um estudo é válido ou ele é inválido, e ou é relevante ou é irrelevante. Uma conclusão baseada em estudos inválidos ou irrelevantes não pode ser relevante ou confiável e, por essa razão, deve ser rejeitada").

qualquer dos seus componentes isoladamente – podendo ser, na terminologia de *Daubert*, mais confiável do que qualquer dos seus componentes.

Pode-se pensar – e por um tempo eu mesma pensei – que essa dificuldade poderia ser evitada se *Daubert* fosse interpretado como exigindo, não que cada item do testemunho especializado aponte de modo suficientemente confiável à conclusão final de que exposição a S causa ou fomenta D, mas que cada item aponte de modo suficientemente confiável a conclusão do estudo ao qual se faz referência: por exemplo, que os dados de um estudo epidemiológico apontem de modo suficientemente confiável a conclusão "há um aumento do risco de $n\%$, entre aqueles que foram expostos a S, de desenvolverem D", que os dados de um estudo com animais apontem de modo suficientemente confiável a conclusão "quando animais desse tipo são expostos a tal dosagem de S, administrada desta maneira, $m\%$ deles desenvolvem D", e assim por diante. Porém, mesmo que possivelmente haja alguma justificação para essa interpretação da sentença na nota de rodapé escrita pelo ministro Blackmun sobre o significado pretendido de confiabilidade probatória¹¹⁰, receio que isso não resolva o problema.

"[A] exatidão ilusória", observou Oliver Wendell Holmes certa vez com perspicácia, "é a fonte de falácias por toda a lei"¹¹¹. E, de fato, não está claro se seria desejável atribuir um significado preciso a "preponderância da prova", mesmo isso fosse possível. Porém, para os propósitos do meu argumento não importa ao que exatamente o padrão da "preponderância da prova" – uma expressão que, de maneira bastante interessante, traz embutida a metáfora do "peso"¹¹² – equivale. Porque o ponto essencial é que, ainda que se institua o padrão, pode haver exemplos em que a prova é equilibrada de modo parêlo, isto é, em que a prova justifica C e não-C no mesmo grau; e, em tais circunstâncias, até mesmo um incremento mínimo da justificação para um lado ou para o outro propiciaria uma "preponderância" a favor de C ou contra C. E, mesmo sendo verdade que a prova p_1 favorável a C, vá incrementar menos a justificação de C se ela própria não chegar a ser sólida (e a prova p_2 , desfavorável a C, vá diminuir menos a justificação de C se ela própria não chegar a ser sólida), até mesmo tal prova pode desequilibrar a balança, isto é, fazer a diferença entre "equilibrado por igual" e "marginalmente favorecendo C frente a não-C" ou vice versa. E assim, se algum elemento de prova que possa ter desequilibrado a balança for excluído com base na cláusula de confiabilidade de *Daubert*, isso pode, na verdade, impedir a ava-

110 *Daubert*, 509 U.S. p. 590 n.9 (caracterizando a "confiabilidade probatória").

111 *Truax v. Corrigan*, 257 U.S. 312, 341 (1921).

112 *Webster's Ninth New Collegiate Dictionary* 929 (Merriam Webster, 1991) (define "preponderância" como "superioridade em termos de peso, poder, importância ou força"). (Compare com "ponderoso").

liação da confiabilidade do testemunho científico em sua totalidade – porque o júri nunca chegará a examinar qualquer elemento que o tribunal exclua em razão de ser insuficiente por si mesmo para atender ao padrão.

É claro que, embora a verdade factual seja inegavelmente importante para a justiça substantiva, algumas regras de prova – por exemplo, segredo marital ou *FRE 407(b)*, sob a qual a prova de conserto subsequente é inadmissível – permitem deliberadamente que considerações de natureza política impeçam a apresentação de prova que pode vir a ser altamente relevante para a verdade dos fatos em pauta. Se tais regras politicamente orientadas são justificáveis é uma questão totalmente distinta, que não posso aprofundar aqui¹¹³; mas, de qualquer modo, a *FRE 702* não é uma regra desse tipo, mas está focada precisa e inequivocamente na confiabilidade.

Mais relevante para o nosso argumento atual é pensar que, ao excluir o testemunho científico com base na cláusula de confiabilidade de *Daubert*, os tribunais podem (ao menos algumas vezes) ter sido motivados pela preocupação de que um júri confrontado com um grande número de provas fracas poderia chegar a uma conclusão não justificada. Um júri de fato pode ser enganado dessa maneira: porque, como argumentei aqui, do fato de que uma combinação de peças de prova individualmente insuficientes pode em seu conjunto justificar uma conclusão num grau mais elevado do que qualquer dos elementos que a compõem não decorre que toda e qualquer combinação dessas provas justifique a conclusão de acordo com o padrão de prova exigido. Porém, um tribunal também pode ser enganado, talvez na direção oposta; porque do fato de que uma combinação de peças de prova individualmente insuficientes pode em seu conjunto justificar uma conclusão num grau mais elevado do que qualquer dos elementos que a compõem não decorre que toda e qualquer combinação dessa prova deixe de justificar a conclusão de acordo com o padrão de prova exigido. Isso revela que a raiz do problema é que, enquanto o sistema legal se baseia de modo crescente no testemunho científico, nem juízes nem júris – nem advogados, nesse tocante – estão bem aparelhados para fazer julgamentos sobre questões científicas nas quais até mesmo especialistas altamente qualificados e competentes possam honesta e razoavelmente estar em desacordo.

113 Ver, porém, “Epistemologia legalizada...”, pp. 151-172 neste volume, 000-000 (tratando brevemente desse assunto).