

Raciocínio sobre probabilidades condicionais: as evidências a favor da hipótese freqüentista se fundamentam em comparações errôneas

David O'Brien

City University of New York
Universidade Federal de Pernambuco

Antonio Roazzi

Maria da Graça B. B. Dias
Universidade Federal de Pernambuco

Resumo

Gigerenzer e Hoffrage (1995) e Cosmides e Tooby (1996) propuseram uma *hipótese freqüentista* a qual afirma que, embora as pessoas raramente façam julgamentos sobre probabilidades condicionais que estejam de acordo com padrões normativos do teorema de Bayes, tais julgamentos podem ser feitos mais comumente quando um problema é apresentado em termos de freqüências do que em probabilidades. Esses dois artigos apresentam 10 experimentos com 89 versões de problemas que apóiam sua afirmação, com sujeitos resolvendo, mais freqüentemente, de forma consistente, problemas no formato freqüentista do que no formato probabilista. Os resultados de dois experimentos relatados no presente estudo mostram, contudo, que os achados destes autores podem ser explicados como resultantes de dois problemas metodológicos que têm relação com a presença versus ausência de um formato de resposta e com o uso de números inteiros *versus* frações decimais. No Experimento 1, encontrou-se que os problemas freqüentistas eram resolvidos apenas quando apresentados com um formato de resposta que pode estimular suposições precisas e, no Experimento 2, que problemas de formato freqüentista e probabilista eram resolvidos na mesma freqüência quando apresentados com formato de resposta e com números inteiros. Os resultados são discutidos em termos das suas implicações contra a hipótese freqüentista.

Palavras-chave: hipótese freqüentista; probabilidades condicionais; teorema de Bayes; problemas de formatos freqüentista e probabilista

Abstract

Reasoning about conditional probabilities: the evidence for the frequentist hypothesis has relied on flawed comparisons. Gigerenzer and Hoffrage (1995) and Cosmides and Tooby (1996) proposed a *frequentist hypothesis* that claims that although people rarely make judgements about conditional probabilities that concord with the normative standards of Bayes's theorem, such judgements can be elicited when a problem is presented in terms of frequencies rather than probabilities. These two articles together reported 10 experiments with 89 problem versions in support of their prediction, with people consistently solving frequentist-formatted problems more frequently than probabilist-formatted problems. The results of two experiments reported here, however, show that their results can be explained as resulting from two experimental confounds that have to do with the presence vs. absence of a response format and with the use of whole numbers vs. decimal fractions. In Experiment 1 we found that the frequentist problems were solved only when presented with a response format that can encourage accurate guesses, and in Experiment 2 we found that frequentist- and probabilist- formatted problems were solved equally often when presented with the response format and with whole numbers rather than with decimal fractions. The results are discussed in terms of their negative implications for the frequentist hypothesis.

Key-words: frequentist hypothesis; conditional probabilities; Bayes's theorem; frequentist- and probabilist- formatted problems

Um debate relativo à natureza das habilidades de raciocínio básico que são úteis para fazer julgamentos sobre probabilidades condicionais tem sido proeminentemente na recente literatura sobre psicologia cognitiva, publicados em revistas como *Psychological Review* e *Cognition*.

O debate começou com uma série de artigos de Gigerenzer e colaboradores (e.g., Gigerenzer, 1991; 1994; Gigerenzer, Hell, & Blank, 1988; Gigerenzer & Hoffrage, 1995; Gigerenzer & Murray, 1987), tendo continuação com Cosmides e Tooby (1996), criticando a afirmação de Kahneman e Tversky (e.g.,

1972, 1973, 1982) de que as pessoas raramente raciocinam de maneira consistente com os tipos de padrões normativos apresentados por teorias formais de probabilidade, mas, ao contrário, tendem a ser guiados por uma variedade de vieses e heurísticas de raciocínio.

As pesquisas de Kahneman e Tversky consistiam amplamente em apresentar problemas de raciocínio para os quais há uma resposta que eles poderiam apontar como normativamente adequada, mostrar que seus sujeitos geralmente não davam essa resposta normativamente adequada e, então, argumentar que a resposta era, ao invés de normativamente adequada, consistente com um ou outro conjunto de heurísticas que eles descreveram (e.g., representatividade, prototipicalidade, ancoragem, disponibilidade).

Uma das principais descobertas que Kahneman e Tversky relataram foi uma falha geral de participantes de pesquisas em resolver problemas que requeriam fazer um julgamento sobre uma probabilidade condicional do tipo que poderia ser computada usando o Teorema de Bayes, com participantes geralmente ignorando a informação da probabilidade de base. Por exemplo, na pesquisa do problema do táxi era dito aos sujeitos que em uma cidade onde apenas 15% dos táxis eram azuis, havia um táxi que tinha se envolvido em um acidente, e este táxi havia sido identificado como azul por uma testemunha que era precisa em fazer tais identificações 80% das vezes. Quando se perguntava aos sujeitos sobre a probabilidade de ser o táxi do acidente, de fato, azul, eles deixavam de considerar que táxis azuis são relativamente incomuns e consideraram, ao invés disso, apenas que a testemunha usualmente fosse precisa. (ver também Bar-Hillel, 1980).

De acordo com Gigerenzer e seus colegas, Kahneman e Tversky cometeram um grave erro conceitual quando afirmaram que problemas como este poderiam ser resolvidos usando o teorema de Bayes¹. O problema conceitual, argumentaram Gigerenzer e seus colegas, é que o teorema de Bayes deveria ser usado apenas para computar frequências populacionais, mas não para computar as probabilidades de eventos únicos, e estes problemas requeriam que fossem feitos julgamentos sobre as probabilidades de eventos únicos². Eles argumentaram que os problemas apresentados por Kahneman e Tversky, assim, não forneceram um teste apropriado em que o teorema de Bayes fosse uma solução normativamente correta, porque o teorema de Bayes não deveria ter sido aplicado a esses tipos de problemas.

Embora a afirmação de que, não se deve julgar as probabilidades de eventos únicos, derive de considerações da literatura sobre teoria da probabilidade, Gigerenzer e colaboradores propuseram um resultado psicológico, alegando que a mente está disposta a representar informações sobre frequências, mas não probabilidades e, conseqüentemente, que a mente está disposta a fazer julgamentos sobre frequências, mas não probabilidades³. A partir desta perspectiva, os sujeitos dos experimentos de Kahneman e Tversky falharam em responder, de uma forma consistente com o teorema de Bayes,

porque eles foram solicitados a resolver problemas sobre as probabilidades de eventos únicos, ao invés de frequências de eventos. Se fosse solicitado aos sujeitos que resolvessem problemas de frequências populacionais, eles deveriam ser capazes de fazê-lo.

Os dois artigos mais influentes que defendem esta hipótese freqüentista (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995) apresentaram grandes quantidades de resultados, objetivando mostrar que julgamentos correspondiam à solução do teorema de Bayes quando os problemas diziam respeito a frequências de população, mas não quando eles diziam respeito a probabilidades de eventos únicos (no total, foram apresentados 17 problemas em oito experimentos de Cosmides & Tooby e 72 problemas foram apresentados em dois experimentos de Gigerenzer & Hoffrage). O foco dos experimentos relatados no presente artigo se refere a: se se deve aceitar estes dados como favoráveis à hipótese freqüentista, como Gigerenzer e Hoffrage, e Cosmides e Tooby argumentaram, ou se seus dados se fundamentam em outras características da tarefa que eram estranhas à comparação de interesse teórico, i.e., formato de problemas freqüentistas *versus* formato de problemas probabilistas. Trataremos aqueles problemas que dizem respeito à probabilidade de um evento único como sendo problemas probabilistas e, aqueles problemas que dizem respeito à frequência de eventos de um tipo particular, como sendo problemas freqüentistas. Para ilustrar porque não se deve aceitar os dados como favoráveis à hipótese freqüentista, considere-se dois dos problemas de Cosmides e Tooby (1996) que em uma primeira impressão parecem apresentar a comparação mais clara entre as questões probabilista e freqüentista (referindo-se aos problemas de Cosmides e Tooby como problemas E5 e E7 - C1). Os dois problemas apresentaram o mesmo parágrafo introdutório:

A prevalência da doença X é 1/1000. Um teste foi desenvolvido para detectar quando uma pessoa tem a doença X. Cada vez que o teste é realizado em uma pessoa que tem a doença, o resultado do teste é positivo. Mas, algumas vezes, o teste também apresenta um resultado positivo quando é realizado em uma pessoa que é completamente saudável. Especificamente, 5% de todas as pessoas que são perfeitamente saudáveis apresentam um resultado positivo para a doença.

Os dois, então, diferiam apenas em termos da questão que era perguntada aos sujeitos.

A versão probabilista apresentava a seguinte questão: “qual é a chance de uma pessoa que teve o resultado positivo ter, de fato, a doença, assumindo que você nada sabe sobre os sintomas ou sinais da pessoa? _____%”

A versão freqüentista apresentava a seguinte questão: “quantas pessoas, cujo teste deu positivo para a doença, terão realmente a doença? ____ de ____.”

Cosmides e Tooby (1996) relataram que a versão freqüentista conduz à resposta Bayesiana muito mais freqüentemente do que a versão probabilista (64% e 36%, respectivamente) o que, à primeira vista, dado que os proble-

mas eram idênticos - exceto pelo fato de um problema apresentar a questão freqüentista e o outro a questão probabilista - parece fornecer forte evidência para suas hipóteses.

Note-se, contudo, que as duas questões diferem em aspectos que escapam à comparação entre os formatos freqüentista e probabilista. Primeiramente, a escrita usada na questão freqüentista parece muito mais objetiva do que a usada na questão probabilista.

De fato, a questão probabilista poderia ter sido apresentada em paralelo à questão freqüentista, como segue: qual é a probabilidade de uma pessoa, cujo teste deu positivo para a doença, ter realmente a doença? ____ de ____ .

Ou a questão freqüentista poderia ter sido apresentada paralelamente às questões probabilistas, como segue: quantas pessoas, cujo teste deu positivo para a doença, terão realmente a doença, considerando que você nada sabe sobre os sintomas ou sinais da pessoa?

Em segundo lugar, as duas questões foram apresentadas com dois formatos de resposta diferentes: enquanto o problema probabilista solicitou aos participantes que respondessem no formato “__ %”, o problema freqüentista solicitou aos participantes que respondessem no formato, “__ de ____”. Note-se que o formato “__ de ____” fornece uma pista de que a solução da tarefa requer estabelecer uma razão e, embora esta pista esteja presente no formato de resposta da versão freqüentista, ela está ausente no formato de resposta fornecido pela versão probabilista.

A diferença no formato de resposta (que aparentemente favorece o formato freqüentista, ao invés do probabilista) poderia ser aceitável se se requisitasse a apresentação de uma comparação entre problemas probabilista e freqüentista. De fato, os dois problemas poderiam ter sido apresentados sem qualquer formato de resposta, apenas com as questões que precedessem estes formatos. Finalmente, através dos oito experimentos, Cosmides e Tooby incluíram o formato de resposta “__ de ____” em cada versão freqüentista e o formato de resposta “__ %” em cada versão probabilista, confundindo, assim, a comparação de crucial interesse (uma possível diferença entre versões freqüentista e probabilista) com uma comparação de nenhum interesse teórico (a diferença entre o formato de resposta “__ de ____” e o formato de resposta “__ %”). Além do mais, Gigerenzer e Hoffrage aparentemente repetiram este modelo em todos os 72 problemas que apresentaram e, assim, as comparações relatadas por Gigerenzer e Hoffrage são igualmente equivocadas⁴. Assim, é de considerável interesse descobrir em que intensidade a inclusão ou exclusão do formato de resposta “__ de ____” pode ser responsável pelas razões de solução aparentemente maiores das versões do problema freqüentista quando comparado ao “__ %” fornecido pelas versões probabilistas.

Experimento 1

Foram apresentados três pares de problemas, todos em formato freqüentista. Entretanto apenas um dos problemas de cada par incluía o formato de resposta “__ de ____”.

Nosso interesse com isto era descobrir se a intensidade de sucesso na resolução desses problemas freqüentistas era dependente da presença do formato de resposta. Foram apresentados os dois problemas freqüentistas extraídos da Tabela 1 de Gigerenzer e Hoffrage, e o problema freqüentista extraído do Experimento 7, de Cosmides e Tooby, anteriormente descrito.

Método

Participantes

Participaram 150 estudantes universitários que foram recrutados nos corredores da Universidade Federal de Pernambuco.

Tarefas e Procedimentos

Foram apresentados três pares de problemas. Um problema de cada par foi apresentado com o formato de resposta “__ de ____”, e o problema acompanhante foi apresentado sem este formato de resposta. Os dois problemas que eram idênticos aos problemas freqüentistas da Tabela 1 de Gigerenzer e Hoffrage eram os seguintes:

Problema 1

Dez de cada 1000 mulheres com 40 anos de idade que participaram de um exame de rotina tinham câncer de mama. Oito de cada 10 mulheres com câncer de mama terão uma mamografia positiva. Noventa e cinco de cada 990 mulheres sem câncer de mama também terão uma mamografia positiva. Aqui está uma nova amostra representativa de mulheres com 40 anos que têm mamografia positiva num exame de rotina. Quantas dessas mulheres você acha que, realmente, têm câncer de mama? ____ de ____.

Problema 2

Cento e três de cada 1000 mulheres de 40 anos têm uma mamografia positiva num exame de rotina. 8 de cada 1000 mulheres de 40 anos que participam de um exame de rotina têm câncer de mama e uma mamografia positiva. Aqui está uma nova amostra representativa de mulheres com 40 anos que têm uma mamografia positiva num exame de rotina. Quantas dessas mulheres você acha que, realmente, têm câncer de mama? “__ de ____”

O Problema 3 foi o problema freqüentista E7 - C1 de Cosmides e Tooby (1996), que foi apresentado acima na introdução.

Para cada participante foi apresentado um único problema, e 25 participantes foram designados para cada problema. As tarefas foram administradas individualmente ou em pequenos grupos (neste último caso não era permitida a troca de informações). Os participantes foram instruídos sobre o fato de que cada problema exigia uma resposta numérica e lhes foi dito que eles deveriam dar a melhor resposta possível, e não necessariamente a primeira resposta que lhes ocorresse. Finalmente, eles foram solicitados a apresentar uma

breve justificativa escrita para mostrar como eles chegaram àquela resposta. Esta explicação foi exigida principalmente para estimular os sujeitos a se engajarem na tarefa.

Resultados e Discussão

As proporções em que os sujeitos deram respostas que foram classificadas como correspondentes à solução do Teorema de Bayes são apresentadas na Tabela 1.⁵

Tabela 1
Proporções de respostas Bayesianas no Experimento I com e sem o formato de resposta “___ de ___”.

Problema	Formato de Resposta	
	Com	Sem
Problema 1 (G&H Frequência Padrão)	0,08	0,00
Problema 2 (G&H Frequência Curta)	0,32	0,04
Problema 3 (Cosmides & Tooby)	0,24	0,00

Nota: G&H se refere a um problema que foi apresentado em Gigerenzer e Hoffrage (1995). O problema de Cosmides e Tooby (1996) pode ser classificado com um critério Bayesiano rígido ou com o critério menos rígido usado por Cosmides e Tooby. As proporções mostradas para o Problema 3 nesta tabela usaram o critério menos rígido. Quando se usa o critério mais rígido, a proporção cai para 0,04 na condição Com Formato de Resposta.

A análise da Tabela 1 revela que a solução do problema exigia a presença do formato de resposta “___ de ___”. Apenas um sujeito, dos 75 que receberam um problema sem este formato de resposta, forneceu uma solução Bayesiana, contra 15 sujeitos que forneceram este tipo de resposta nos problemas com outro formato ($z = 3,70$; $p < 0,001$). Dois dos três problemas que apresentaram o formato de resposta “___ de ___” foram classificados como Bayesianos significativamente mais vezes do que foram suas versões de problemas acompanhantes sem este formato de resposta ($z = 2,58$; $p < 0,005$ e $z = 2,61$; $p < 0,005$, respectivamente, para os Problemas 2 e 3), mas o Problema 1 não conduziu a respostas Bayesianas, mesmo quando apresentado com o formato de resposta “___ de ___”.

Não está claro, entretanto, se os problemas 2 e 3 conduziram a significativas proporções de respostas Bayesianas porque esses problemas permitiram aos sujeitos raciocinarem de forma Bayesiana, nem se cada problema racionalmente poderia levar a tais respostas sem uma linha de raciocínio Bayesiano. Considere-se primeiro o Problema 2, que apresenta três números (8, 103 e 1000) e um formato de resposta com espaço para dois números (___ de ___). Um sujeito que não tenha idéia sobre como construir qualquer linha de raciocínio para esse problema, poderia decidir colocar dois quaisquer dos três números nos espaços fornecidos. Admitindo que a maioria dos adultos entendeu que o formato de resposta “___ de ___” indica que o número à esquerda deveria ser menor do que o número à direita (i. e., que a relação de inclusão de classe impede a possibilidade de um subconjunto ser maior

que seu conjunto), tal estratégia levaria a três respostas possíveis (8 de 103; 8 de 1000; 103 de 1000), e de fato, essas três categorias de respostas conduziram a 84% do total para esse problema.

Além disso, se os sujeitos estavam seguindo tal estratégia, seria esperado, no Problema 2, que um em cada três sujeitos escolhesse a resposta 8 de 103, que é quase exatamente a proporção mostrada na Tabela 1 para essa resposta nesse problema. Assim, a resposta Bayesiana dada por 32% dos sujeitos no Problema 2 não foi seguida de uma linha de raciocínio Bayesiano, mas, ao invés disto, de uma simples estratégia de preencher os espaços de formato de resposta com dois quaisquer dos três números apresentados no problema. Denominamos esse modo de obter a resposta Bayesiana de “estratégia sorte de tolo”, e o uso de tal estratégia também explicaria porque as respostas Bayesianas foram raramente dadas no Problema 1. Note-se que tal estratégia não aparece no Problema 1 como um caminho que conduz à resposta “8 de 103”; para isso exige que dois números (8 + 95) precisem ser somados para preencher o lado direito do formato de resposta apropriadamente. Desse modo, uma estratégia que preenche meramente os espaços com os números apresentados no problema poderia não conduzir a respostas Bayesianas no Problema 1.

Considere-se agora o Problema 3 (Problema E-7 C1, de Cosmides & Tooby, 1996). Cosmides e Tooby relataram que 64% de seus sujeitos deram respostas Bayesianas nesse problema, mas, como visto acima, os autores usaram um critério um tanto generoso, e classificaram como Bayesianas respostas que não seriam contadas se usassem um critério de classificação mais rígido. Consideremos o que seria uma linha de raciocínio Bayesiano nesse problema. Precisar-se-ia computar a razão entre aquelas pessoas que seriam identificadas corretamente como portadoras de uma doença e todas aquelas identificadas como portadoras da doença, incluindo aquelas corretamente identificadas, bem como aquelas incorretamente identificadas. Usando o número de pessoas apresentado quando o problema afirma que a prevalência da doença é 1/1000 e sabendo que o teste sempre identifica corretamente que uma pessoa tem a doença, pode-se facilmente preencher à esquerda do formato de resposta como igual a 1,00. Isso deixa 999 pessoas livres da doença, das quais 5% estarão falsamente identificadas como portadoras da doença (= 49,95 pessoas). Precisar-se-ia, agora, adicionar estas 49,95 pessoas falsamente identificadas àquela pessoa corretamente identificada para computar o número total de pessoas identificadas (= 50,95) e assim preencher o lado direito do formato de resposta. A resposta Bayesiana, então, é 1 de 50,95, embora Cosmides e Tooby contassem 1 de 50 como Bayesiana. Esse é também o resultado apresentado como Bayesiano na Tabela 1, e é válido perguntar se isso poderia ter inflacionado injustamente suas estimativas de respostas Bayesianas para esse problema.

Consideremos se seria aceito um método de classificação similar quando confrontado com um problema com diferentes valores numéricos, tal como o problema do táxi des-

critério anteriormente, com sua probabilidade de base = 15 para táxis azuis e com uma testemunha que identifica precisamente = 80. Nesse caso, a solução Bayesiana seria 12 de 12+17 (o número de táxis que seria corretamente identificado dividido pelo número total de táxis que seria identificado como azul pela testemunha, incluindo aqueles identificados corretamente e aqueles falsamente identificados). Nesse problema, entretanto, a resposta que corresponde, computacionalmente, à resposta “1 de 50”, que foi classificada por Cosmides e Tooby como Bayesiana em seu problema E-7 C-1, seria “12 de 17”, que é extremamente diferente de “12 de 29” (i.e. 71 ao invés de 41) que é o resultado do teorema Bayesiano. A semelhança da resposta Bayesiana (1 de 50,95) e a resposta classificada por Cosmides e Tooby como Bayesiana (1 de 50) é meramente uma função dos números apresentados em seu problema e é preciso não aceitar que a resposta “1 de 50” indique necessariamente uma linha de raciocínio Bayesiano. Pode-se pensar que 1 de 50 seria aceito como arredondamento (afinal, não podem existir 49,95 pessoas). Embora, o arredondamento possa, por si mesmo, ser aceitável por um adepto de uma linha de raciocínio correspondente ao teorema de Bayes, ainda assim não temos razão para pensar que classificaríamos 1 de 50 como a resposta arredondada correta, porque a resposta que resultaria do arredondamento seria 1 de 51.

O que isso tem a ver com nossa discussão anterior do formato de resposta “___de ___” e com a suposição de que um sujeito, que seguiu a estratégia de preencher meramente os espaços em branco do formato de resposta com números que foram apresentados no problema, poderia resolver o problema pela sorte de tolo? Essa mesma estratégia de sorte de tolo pode ser aplicada ao Problema 3? A resposta é “sim”, com apenas algumas modificações. Uma pessoa que raciocina equipada com tal estratégia e confrontada com o Problema 3, precisaria simplesmente computar que 5% de 1000 é 50 para preencher ao lado direito do formato de resposta e, então, ser classificada, conforme o critério de Cosmides e Tooby, como tendo dado uma resposta Bayesiana, mesmo que essa pessoa que raciocinou não tenha seguido nada equivalente a uma linha de raciocínio Bayesiano. Realmente, quando o formato de resposta “___de ___” estava ausente, nenhum participante apresentou uma solução Bayesiana, que é exatamente a situação com que nos deparamos no Problema 2. Além disso, quando usamos um critério mais rígido para classificar este problema, aceitando como Bayesianas apenas aquelas respostas que eram claramente Bayesianas na inspeção das explicações escritas (1 de 51, ou 1 de 50,95), somente uma das setes respostas que se encontram na Tabela 1 como Bayesianas no Problema 3 com formato de resposta poderia ainda ser classificada como resposta Bayesiana.

Os dados apresentados por Cosmides e Tooby podem, portanto, ter superestimado a probabilidade com que tais problemas conduziram a uma solução Bayesiana.⁶

Em suma, os dados do presente experimento não fornecem qualquer evidência a favor da hipótese freqüentista. Ao invés disto, sugerem que a evidência que havia sido relatada a favor da hipótese freqüentista nos experimentos de

Cosmides e Tooby e de Gigerenzer e Hoffrage, se fundamentam completamente na presença do formato de resposta “___de ___”. Levando em consideração que este formato de resposta estava presente nos problemas freqüentistas que eles apresentaram, mas ausente nos probabilistas, os únicos dados a favor da hipótese freqüentista agora encontram-se sob suspeita de terem sido apenas um artefato de um projeto de pesquisa internamente inválido. Além disso, dado o formato de resposta “___de ___”, é possível considerar uma explicação alternativa para as respostas Bayesianas que realmente apareceram nos problemas 2 e 3: nossa hipótese de sorte de tolo assegura que os sujeitos talvez não tenham seguido uma linha de raciocínio Bayesiano em nenhum momento, mas, ao invés disto, talvez tenham simplesmente preenchido os espaços em branco fornecidos no formato de resposta, utilizando-se dos números dados nos problemas; uma estratégia que poderia conduzir um terço das respostas a imitar as respostas Bayesianas, e os dados se encaixaram perfeitamente à hipótese sorte de tolo. As proporções das respostas Bayesianas nos problemas 2 e 3 quando apresentados com formato de resposta não diferiram da proporção esperada em um evento único (os escores z eram menores que 1,00), mas respostas Bayesianas se mostraram em menor número, diante do que era esperado em um evento único nos outros quatro problemas (os escores z variavam de 2,66 a 3,51). Portanto, a adivinhação fornece uma melhor explicação para os presentes resultados do que as próprias explicações alternativas.

Experimento 2

Uma avaliação dos problemas apresentados na Tabela 1 de Gigerenzer e Hoffrage (1995) – os únicos problemas que apresentaram integralmente – revela outro aspecto da tarefa que pode tornar os problemas freqüentistas mais fáceis do que os probabilistas com os quais foram comparados: os problemas freqüentistas – problemas 1 e 2 em nosso Experimento 1 – apresentaram informações numéricas de números inteiros, ao passo que os problemas probabilistas acompanhantes incluíam informação numérica apresentada em frações decimais. Considere-se o problema 2 do Experimento 1 – ao qual Gigerenzer e Hoffrage se referiam como tendo formatos freqüentistas curtos – que afirmava que 103 de 1000 mulheres apresentaram uma mamografia positiva e 8 de 1000 mulheres apresentaram uma mamografia positiva e um câncer de mama. O problema acompanhante com o formato probabilista curto afirmou que a probabilidade de uma mulher ter uma mamografia positiva é 10,3% e que a probabilidade de ter câncer de mama e uma mamografia positiva é 0,8%. Diferentemente da questão no Experimento 1, relativa à presença versus ausência do formato de resposta do “___de ___”, o fato de ter esse problema freqüentista apresentado números inteiros e seus problemas probabilistas acompanhantes apresentado frações decimais, pode ter influenciado na resolução. A mesma situação se repetiu no problema 1 do Experimento 1 – ao qual Gigerenzer e Hoffrage se referiram como tendo um formato de

freqüência padrão – e seu problema acompanhante, o formato probabilista padrão. Esta possibilidade, que a apresentação diferente de informação numérica com números inteiros para os problemas freqüentistas e com frações decimais para problemas probabilistas, pode ter contribuído para probabilidades diferentes com os quais os dois tipos de problemas foram resolvidos, que nos referimos no Experimento.

Os problemas apresentados aqui diferem dos problemas de mamografia apresentados no Experimento 1 (e daqueles na Tabela 1, de Gigerenzer e Hoffrage), pois eles se referiam não a um teste médico e uma doença, mas, ao contrário, a cartões de formas e cores arbitrárias em um jarro. Esta mudança no conteúdo foi motivada por nossos achados durante a observação de que alguns sujeitos no Experimento 1 se basearam em crenças prévias sobre câncer de mama e mamografia na avaliação da evidência apresentada nos problemas, ao invés de se basearem nas informações numéricas apresentadas.

Método

Participantes

Participaram 200 estudantes universitários que foram recrutados nos corredores da Universidade Federal de Pernambuco.

Tarefas e Procedimentos

Foram apresentados 8 problemas, todos com o formato de resposta “___de ___”. Os procedimentos foram os mesmos utilizados no Experimento 1, novamente cada participante recebendo um único problema, e cada problema sendo designado a 25 participantes. Cada problema referia-se a formas arbitrárias (ou triângulos ou círculos) de duas cores (preto e branco). As formas dos problemas seguiram os quatro tipos de problemas apresentados por Gigerenzer e Hoffrage: probabilista curto, freqüentista curto, probabilista padrão e freqüentista padrão. Em quatro dos problemas, os valores numéricos foram idênticos aos apresentados nos problemas de mamografia de Gigerenzer e Hoffrage; nos outros quatro problemas os valores numéricos foram alterados para que as versões probabilistas fossem apresentadas com números inteiros e as freqüentistas com frações decimais. Estes quatro novos problemas foram formulados de modo que os valores numéricos das versões acompanhantes probabilistas e freqüentistas fossem aritmeticamente equivalentes na mesma forma que os outros quatro problemas apresentados por Gigerenzer e Hoffrage; e.g., 5,95 de 17 = 35%. No início de cada um dos oito problemas, pedia-se aos sujeitos que imaginassem um jarro contendo quatro tipos de cartões (triângulos brancos, triângulos pretos, círculos brancos, círculos pretos), e todos os oito problemas foram organizados da mesma maneira. As informações numéricas e as perguntas para cada um dos oito problemas se encontram na Tabela 2.

Resultados e Discussão

A Tabela 3 mostra as proporções com as quais cada problema levou a respostas consistentes com a solução do teorema de Bayes.

Observa-se que apenas dois problemas foram resolvidos mais freqüentemente do que os outros, ambos foram versões de problemas curtos apresentados com números inteiros, i.e, a versão probabilista curta apresentada com os novos números e a versão probabilista curta apresentadas com os números dos problemas de mamografia. O problema freqüentista curto com números da mamografia (i.e., com números inteiros) conduziu a respostas Bayesianas com mais freqüência do que os problemas freqüentistas curtos com os novos números, $z = 2,09$; $p < 0,05$, e os problemas probabilistas curtos com os números novos conduziram a respostas Bayesianas com mais freqüência do que os problemas probabilistas curtos com os números da mamografia, $z = 1,99$; $p < 0,05$.

Embora tenha havido diferença significativa entre os dois problemas curtos apresentados com números da mamografia, $z = 2,77$; $p < 0,01$ (com mais respostas Bayesianas na versão de formato freqüentista), a diferença entre problemas freqüentistas curtos e probabilistas curtos não foi significativa quando os problemas foram apresentados com novos números. Além disso, entre os quatro problemas padrões também não houve diferenças significativas. Assim, os problemas só foram resolvidos quando apresentados na forma curta, e esses problemas foram resolvidos mais freqüentemente quando apresentados com números inteiros do que com frações decimais.

Juntamente com os resultados do Experimento 1, esses dados revelam que o fato de um problema levar a respostas Bayesianas não tem qualquer relação com o fato de ser o problema apresentado no formato probabilista ou freqüentista, mas sim, primeiramente, com o fato de serem as respostas apresentadas no formato “___de ___”, e, em menor extensão, com o fato de apresentarem números inteiros ou frações decimais.

Esses dados sugerem que os resultados relatados anteriormente a favor da hipótese freqüentista, de fato, dependiam, no mínimo, destas duas características estranhas da tarefa, não tendo relação com qualquer diferença entre as versões probabilistas e freqüentistas.

Discussão Geral

Nosso artigo não é o primeiro a questionar a afirmação de que há um benefício em receber as informações dos problemas em forma de freqüência ao invés de probabilidade. Lewis e Keren (1999), Mellers e McGraw (1999), e Vranas (2000), todos têm fornecido argumentos recentes contra as afirmações de Gigerenzer e Hoffrage (1995) e Cosmides e Tooby (1996). No entanto, Gigerenzer e Hoffrage (1999) responderam a estas recentes críticas da seguinte forma: “*There seems to be a consensus about the effect's existence. The debate has now shifted to the questions of why the effect occurs and what its boundary conditions are*” (p. 425). Entretanto, este tipo de resposta de uma mudança de uma questão sobre o “porquê” para uma questão sobre o “quando” não seria adequado quando deparado com os dados apresentados aqui,

Tabela 2

Números e questões para cada um dos oito problemas apresentados no Experimento 2

<i>Freqüentista Curta com Novos Números</i>	
5,95 de cada 17 cartões são pretos.	
3,40 de cada 17 cartões são ambos triângulos e pretos.	
Em uma amostra randômica de cartões pretos, quantos você espera que sejam triângulos? ____ de ____	
<i>Freqüentista Curta com Números da Mamografia</i>	
103 de cada 1000 cartões são pretos.	
8 de cada 1000 cartões são ambos triângulos e pretos.	
Em uma amostra randômica de cartões pretos, quantos você espera que sejam triângulos? ____ de ____	
<i>Freqüentista Padrão com Novos Números</i>	
4,25 de cada 17 cartões são triângulos.	
3,40 de cada 4,25 triângulos são pretos.	
2,55 de cada 12,75 círculos são pretos.	
Em uma amostra randômica de cartões pretos, quantos você espera que sejam triângulos? ____ de ____	
<i>Freqüentista Padrão com Números da Mamografia</i>	
10 de cada 1,000 cartões são triângulos.	
8 de 10 triângulos são pretos.	
95 de cada 990 círculos são pretos.	
Em uma amostra randômica de cartões pretos, quantos você espera que sejam triângulos? ____ de ____	
<i>Probabilista Curta com Novos Números</i>	
A probabilidade de que um cartão seja preto é 35%.	
A probabilidade de que um cartão seja preto e triângulo é 20%.	
Um cartão escolhido randomicamente é preto. Qual a probabilidade de que seja triângulo? ____ de ____	
<i>Probabilista Curta com Números da Mamografia</i>	
A probabilidade de que um cartão seja preto é 10,3%.	
A probabilidade de que um cartão seja preto e triângulo é 0,8%.	
Um cartão escolhido randomicamente é preto. Qual a probabilidade de que seja um triângulo? ____ de ____	
<i>Probabilista Padrão com Novos Números</i>	
A probabilidade de que um cartão seja um triângulo é 25%.	
Se um cartão é um triângulo, a probabilidade de que seja preto é 80%.	
Se um cartão é um círculo, a probabilidade de que seja preto é 20%.	
Um cartão escolhido randomicamente é preto. Qual a probabilidade de que seja um triângulo? ____ de ____	
<i>Probabilista Padrão com Números da Mamografia</i>	
A probabilidade de que um cartão seja um triângulo é 1%.	
Se um cartão é um triângulo, a probabilidade de que seja preto é 80%.	
Se um cartão é um círculo, a probabilidade de que seja preto é 9,6%.	
Um cartão escolhido randomicamente é preto. Qual a probabilidade de que seja um triângulo? ____ de ____	

Tabela 3

Proporções das respostas Bayesianas com números inteiros e com frações decimais no Experimento 2

Problema	Apresentação numérica	
	Números novos	Números da mamografia
Freqüentista Curta	0,20	0,48
Probabilista Curta	0,36	0,12
Freqüentista Padrão	0,04	0,12
Probabilista Padrão	0,08	0,00

os quais não mostram a existência de simples limites no tipo de características entre as condições; pelo contrário, sugerem que a evidência mais forte que foi apresentada a favor da hipótese frequentista tenha resultado de problemas metodológicos sérios nas comparações experimentais; quando estes problemas foram retirados no presente experimento, a evidência a favor da hipótese frequentista desapareceu.

Os resultados do Experimento 2 são particularmente problemáticos para a afirmação de que o debate diz respeito apenas ao tipo de condições que determinam quando ocorrem os efeitos benéficos da apresentação da informação dos problemas em frequência. As mesmas características da tarefa que impossibilitaram as soluções de problemas nas versões frequentistas também impossibilitaram a solução dos problemas nas versões probabilistas; quando os problemas probabilistas foram apresentados com o formato de resposta “___de ___”, e com números inteiros, ao invés de frações decimais, suas probabilidades de respostas Bayesianas não se diferenciaram significativamente daquelas encontradas nos problemas de formato frequentistas comparáveis, e este é um resultado que a hipótese frequentista ainda não consegue explicar.

Os achados deste experimento não apenas desafiam a afirmação de que é a apresentação da informação do problema em forma de frequências que conduz a proporções elevadas de respostas Bayesianas, mas também levantam a possibilidade de que estes tipos de respostas resultam, não do uso da linha de raciocínio Bayesiano, mas de uma estratégia de adivinhação que deu certo. A descoberta de que a adição do formato de resposta “___ de ___” conduz ao aumento das respostas Bayesianas apenas quando usado em conjunção com a forma curta, levou-nos a propor a hipótese “sorte de tolo”, que assegura que o uso do formato de resposta “___de ___”, juntamente com a forma curta, conduziria um terço dos sujeitos à resposta correta, sem que estes tivessem um *insight* em relação à estrutura do problema. Esse tipo de adivinhação (sorte de tolo), aparentemente pode ser estimulado quando a informação do problema é dada em números inteiros, e pode ser desestimulada quando estas informações são dadas em frações decimais. Não estamos afirmando que a evidência é conclusiva a favor da explicação sorte de tolo, mas essa explicação parece mais plausível como uma razão dos presentes dados do que qualquer um dos outros competidores no debate. Aqueles problemas que apresentaram a forma curta juntamente com o formato de resposta “___de ___” levaram a uma resposta Bayesiana em uma probabilidade que não diferia significativamente de uma probabilidade que seria predita pela hipótese de sorte de tolo ($= 0,33$), enquanto que a probabilidade de tais respostas estava significativamente abaixo das possibilidades nos problemas que não apresentavam tal estratégia. Note-se, em particular, que este achado foi obtido tanto se o problema apresentava frequência ou probabilidade.

Nossa hipótese de sorte de tolo oferece uma explicação perfeita para os resultados destes experimentos, mas ela não fornece uma explicação igualmente perfeita dos dados apresentados por Gigerenzer e Hoffrage (1995) e Cosmides e Tooby (1996), ambos os quais relataram probabilidades de

soluções Bayesianas maiores do que encontramos. Nesse ponto, nossos dados são similares àqueles relatados por outros, e.g., Mellers e McGraw (1999). De fato, Mellers e McGraw sugeriram que as diferenças entre os seus dados e os de Gigerenzer e Hoffrage podem ter ocorrido porque, enquanto no estudo de Mellers e McGraw cada participante recebia apenas um problema (como no nosso experimento), no estudo de Gigerenzer e Hoffrage, os participantes da pesquisa receberam 60 problemas cada. Gigerenzer e Hoffrage (1999) responderam que seus sujeitos deram mais respostas Bayesianas no primeiro problema que receberam do que nos problemas subsequentes, e eles especularam, ao contrário, que a diferença na probabilidade da solução é atribuída a diferenças das populações examinadas, i.e., estudantes universitários americanos *versus* estudantes universitários alemães. Isso não explicaria, certamente, por que os participantes dos experimentos relatados por Cosmides e Tooby (1996) também apresentaram mais respostas Bayesianas do que participantes em outros experimentos relatados. Entretanto, as altas probabilidades das respostas Bayesianas relatadas por Gigerenzer e Hoffrage, e por Cosmides e Tooby, podem ser adequadamente explicadas pela teoria de sorte de tolo. Uma avaliação final sobre quais tipos de explicações seriam mais adequadas para uma melhor compreensão destes resultados precisa aguardar a eventual descoberta de quais estimativas das distribuições populacionais para soluções Bayesianas são mais precisas.

Entretanto, mesmo sem esta consideração, nossos resultados apresentam um sério desafio para aqueles que propõem a hipótese frequentista, dada a nossa descoberta de que aquelas características de tarefas, cruciais para as respostas Bayesianas em problema frequentista, também levam a respostas Bayesianas em problemas probabilistas.

Agradecimentos

Os autores agradecem ao CNPq e ao NSF pelo financiamento da pesquisa que deu origem a este artigo: Convênio de Cooperação Bilateral CNPq (Proc. número 910023/01-8) / NSF (*National Science Foundation*, EUA), iniciado a partir de 1º de julho de 2001, envolvendo auxílios financeiros tanto nos EUA como no Brasil para realização do Projeto de Pesquisa Conjunta “Perspectiva transculturais e evolutivas na linguagem do pensamento”. O projeto é coordenado no Brasil pelo Professores Antonio Roazzi e Maria da Graça Dias e nos EUA pelos Professores David P. O'Brien e Patricia Brooks. No Brasil, o Convênio é gerenciado pela Assessoria de Cooperação Internacional do Conselho Nacional de Desenvolvimento Científico e Tecnológico (ASCIN/CNPq). O número do processo no NSF é INT-0104503, *U.S.-Brazil Cooperative Research: “Cross Cultural and Cross Linguistic Investigations of Deduction and Quantification”*.

Referências

- Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, 44, 211-233.

- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58, 1-73.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases." *European Review of Social Psychology*, 2, 83-115.
- Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is relevant for psychology (and vice versa). In G. Wright & P. Ayton (Orgs.), *Subjective probability* (pp. 129-162). Nova York: Wiley.
- Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 513-525.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684-704.
- Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, New Jersey: Erlbaum.
- Gigerenzer, G., & Hoffrage, U. (1999). Overcoming difficulties in Bayesian reasoning: A reply to Lewis and Keren (1999) and Mellers and McGraw (1999). *Psychological Review*, 106, 425-430.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Orgs.), *Judgment under uncertainty: Heuristics and biases* (pp. 509-520). Cambridge, Reino Unido: Cambridge University Press.
- Lewis, C., & Keren, G. (1999). On the difficulties underlying Bayesian reasoning: A comment on Gigerenzer and Hoffrage. *Psychological Review*, 106, 411-416.
- Mellers, B. A., & McGraw, A. P. (1999). How to improve Bayesian reasoning: Comment on Gigerenzer and Hoffrage (1995). *Psychological Review*, 106, 417-424.
- Vranas, P. B. M. (2000). Gigerenzer's normative critique of Kahneman and Tversky. *Cognition*, 76, 179-193.

Nota

- ¹ No problema do táxi, o teorema de Bayes computa a razão entre a probabilidade de que um táxi seja corretamente identificado como azul e a probabilidade total de ser identificado com azul, incluindo as probabilidades das identificações verdadeira e falsa.
- ² Não nos posicionamos a favor do argumento feito por Gigerenzer e colaboradores de que eventos únicos não devem ser pensados como tendo probabilidades.
- ³ Gigerenzer e Hoffrage (1999) têm enfatizado que suas afirmações dizem respeito apenas a *freqüências naturais* oriundas de amostras naturais e não *freqüências* oriundas completamente de amostras sistemáticas das quais probabilidades de base existem anteriormente a observações. Esta distinção não é importante em nenhuma maneira óbvia para os materiais apresentados aqui.
- ⁴ Gigerenzer e Hoffrage apresentaram apenas um subconjunto de seus 72 problemas na forma completa (na Tabela 1), e estes problemas sempre incluíam o formato "___de ___" em versões freqüentistas e o formato "___%" em versões probabilistas.
- ⁵ Nos problemas 1 e 2, a resposta correta era considerada "8 de 103". No Problema 3 seguimos Cosmides e Tooby em aceitar as respostas "1 de 50" bem como "1 de 51" como corretas, embora, como discutimos, este critério de classificação é altamente liberal.
- ⁶ Estamos sem nenhuma explicação para o fato de que Cosmides e Tooby relataram que 64% das respostas foram Bayesianas neste problema, enquanto apenas 24% o foram no presente experimento, mesmo com o critério de classificação menos rígido que eles usaram.

David O'Brien, doutor em Psicologia pela Temple University (Filadélfia, Pensilvânia), é professor no Departamento de Psicologia e Graduate Center, City University of New York. E-mail: david_obrien@baruch.cuny.edu
Antonio Roazzi, doutor em Psicologia pela University of Oxford (Reino Unido), é professor no Departamento de Psicologia, Universidade Federal de Pernambuco. E-mail: roazzi@npd.ufpe.br

Maria das Graças Bompastor Borges Dias, doutora em Psicologia do Desenvolvimento pela University of Oxford (Reino Unido), é professora no Departamento de Psicologia, Universidade Federal de Pernambuco. E-mail: mdias@npd.ufpe.br

Endereço para correspondência: Universidade Federal de Pernambuco, CFCH, 8º andar; Cidade Universitária; Rua Acadêmico Hélio Ramos, s/n; Recife, PE; CEP 50670-901. Fone: (81) 3271-8272 e 3271-0599. Fax: (81) 3271-1843.