

ANTECIPAÇÃO ANÔMALA DE EVENTOS FUTUROS E ALEATÓRIOS

Em 2011, o psicólogo e professor emérito da *Cornell University*, Daryl J. Bem, conseguiu ver publicado, sob o procedimento de par-revisão, nove experimentos do tipo *pressentimento* numa prestigiosa revista de psicologia (*Journal of Personality and Social Psychology*, [vol. 100, n.º. 3, 407–425](#)), o que proporcionou vasta discussão acadêmica com forte repercussão na imprensa do mundo inteiro (veja aqui algumas reportagens nacionais a respeito: [01](#); [02](#); [03](#)). Tudo isso ocorreu porque os resultados descritos no artigo foram altamente significativos a favor da existência de psi, especificamente sobre a hipótese de que um evento futuro poderia afetar nossas respostas (afetivas ou cognitivas) no presente (numa espécie de efeito *retrocausal*). Nos parágrafos a seguir vamos analisar o debate que cercou tais experimentos.

Para começo, vou sumarizar os experimentos efetivados por Bem na *Cornell University*, mas por questão de tempo abrirei mão de detalhar as variações de alguns deles. Por exemplo, os experimentos de *pressentimento* que abordaram os efeitos psicológicos da *habituação*, da *facilitação de memória* e do [priming afetivo](#) possuíram duas variações cada um. Neles me aterei somente ao Tipo I.

Os experimentos de Bem podem ser vistos como descendentes direto dos testes de *pressentimento*. Como estes, cada um de seus experimentos modificou um efeito psicológico já bem estabelecido, invertendo a usual sequencia temporal dos eventos para que as respostas do participante fossem obtidas antes de os eventos-estímulos presumidamente causais ocorressem. A hipótese em cada um dos casos era a de que a versão de tempo-invertida do experimento produziria o mesmo resultado do experimento padrão de tempo não-invertido. Quatro efeitos psicológicos bem estabelecidos foram assim adaptados: o de aproximação/fuga; o *priming* afetivo, a *habituação* e a *facilitação de memória* (Bem, Tressoldi, Rabeyron e Duggan, 2014).

Em todos esses estudos (com exceção do último experimento - de *facilitação retroativa da memória*) os participantes foram obrigados a fazer rápidos julgamentos em cada tentativa sem tempo para refletir sobre suas decisões. A sequência de eventos dentro de cada julgamento ocorria numa escala de tempo de *milissegundos* e o estímulo presumidamente causal aparecia imediatamente após cada uma das respostas do participante. “Nos termos da teoria da cognição de Kahneman, a do *modo duplo de processamento* (2011) — conforme descrito em seu livro, *Thinking, Fast and Slow* — estas experiências requerem processos cognitivos característicos do sistema 1, de '*pensamento rápido*' (ver também Evans, 2008, e Evans e Stanovich, 2013). Em contraste, o protocolo do experimento de *facilitação retroativa da memória* confrontou os participantes com uma única tarefa cognitiva que se estendida numa escala de tempo de minutos: apresentação de uma lista inicial de palavras que levou de 2-1/2 minutos; o teste de memória de até 5 minutos; e os exercícios práticos pós-teste que levaram aproximadamente 7 minutos” (Bem, Tressoldi, Rabeyron e Duggan, 2014).

Debate Psi
Uma Compilação da Evidência [Ano I, 2014]
www.debatepsi.com

Experimentos com base no sistema de ‘pensamento rápido’

- A detecção precognitiva de um estímulo erótico. 100 estudantes universitários, 50 mulheres e 50 homens, foram recrutados para essa experiência. Os participantes foram avisados de que se tratava de um experimento tipo ESP e de que tinha que adivinhar em qual de duas cortinas virtuais (exibidas lado a lado na tela de um monitor) seria exibida uma imagem. Foi dito a cada participante que teria 36 tentativas e que muitas das imagens continham cenas eróticas. Cada sessão possuía fotografias eróticas e não-eróticas randomicamente misturadas e a *hipótese psi* era de que os participantes seriam capazes de identificar a posição da imagem erótica (i.e., em qual das cortinas ela estaria) mais significativamente do que os 50% esperados pelo acaso. Do ponto de vista dos participantes o procedimento seria parecido com um teste de clarividência, todavia, nem o tipo de foto (erótica ou não erótica) e nem sua localização (cortina da esquerda ou da direita) eram determinadas até o momento de o participante registrar sua escolha. De fato, somente *após* a escolha do participante algum processo randômico iria determinar o tipo de imagem e sua localização.

Através das 100 sessões, os participantes identificaram corretamente a futura posição das imagens eróticas significativamente mais do que seria esperado pelo acaso (53,1% contra 50%), $t(99) = 2.51$, $p = 0.01$, *tamanho do efeito (es) = 0.25*. Não houve diferença significativa entre a taxa de acerto de homens e a de mulheres.

- A fuga precognitiva de um estímulo negativo. 150 universitários participarem desse experimento (107 mulheres e 43 homens). Eles foram informados que o estudo se tratava de um teste ESP. Em cada tentativa do experimento (do total de 36) seria exibida uma imagem e sua respectiva imagem-espelho lado a lado; e então seria solicitado ao participante para indicar qual das imagens ele gostou mais. Foi dito também que o computador iria exibir um *flash* de uma imagem escondida sobre a tela. Nesse experimento, o participante não estava encarando explicitamente um desafio de ESP. Sua tarefa era apenas expressar, em cada tentativa, a sua preferência entre 2 imagens neutras (i.e., de baixa estimulação emocional) e intimamente combinadas. Somente após o participante registrar sua escolha (qual das duas imagens ele gostou mais), o computador iria selecionar aleatoriamente uma das imagens para ser o alvo. Caso o participante tivesse escolhido a imagem não-alvo seria contado como um acerto. Se tivesse escolhido a imagem alvo, lhe seria subliminarmente exibido (por 33ms) uma outra fotografia, mas agora emocionalmente negativa. A *hipótese psi* seria a de que os universitários iriam evitar o estímulo negativo do futuro, preferindo a imagem não-alvo significativamente mais do que os 50% esperados pelo acaso.

De fato, o experimento confirmou a hipótese psi, com os universitários escorando significativamente mais do que os 50% esperados pelo acaso, ainda que o *tamanho do efeito* tenha sido menor do que o primeiro experimento: 51,7% contra 50% esperáveis pelo acaso, $t(149) = 2.39$, $p = 0.009$ e *tamanho do efeito (es) = 0.20*.

- Priming retroativo. Num típico experimento de *priming afetivo*, os participantes são convidados a julgar tão rapidamente quanto podem saber se uma foto lhe é agradável ou desagradável, sendo o tempo de resposta medido. Pouco antes de a imagem aparecer, uma palavra positiva ou negativa (por exemplo, belo, feio) é exibida

Debate Psi

Uma Compilação da Evidência [Ano I, 2014]

www.debatepsi.com

por alguns instantes na tela; essa palavra é chamada *prime*. As pessoas geralmente respondem mais rapidamente quando as valências do *prime* e a imagem são congruentes (ambas são positivas ou ambas são negativas) do que quando elas são incongruentes entre si. Em nossa versão com efeitos retroativos do procedimento, o *prime* apareceu depois, e não antes de os participantes realizarem seus julgamentos das imagens. Foram utilizados 100 universitários, 69 mulheres e 31 homens, com 64 tentativas para cada. A hipótese *psi* era de que o procedimento retroativo também produziria respostas mais rápidas sobre tentativas congruentes do que em relação a tentativas incongruentes.

Num teste de *priming* padrão (sem retrocausalidade) o procedimento alcançou o resultado habitual. Por exemplo, com um critério de corte de 1.500-ms e de transformação inversa, os participantes foram 23,6ms mais rápidos sobre tentativas congruentes do que incongruentes, $t(96) = 4,91$, $p < 0.00001$, *tamanho do efeito* (es) = 0.45. Com o procedimento retrocausal aconteceu o mesmo, predizendo o efeito *psi*. Com um critério de corte de 1.500-ms e de transformação inversa, os participantes foram 15,0ms mais rápidos sobre tentativas congruentes do que incongruentes, $t(96) = 2,55$, $p = 0.006$ e *tamanho do efeito* (es) = 0.25.

- A *habituação retroativa*. Quando indivíduos são expostos inicialmente a um estímulo surpreendente ou emocional, de regra, eles têm uma forte resposta fisiológica. Após repetidas exposições, essa resposta diminui. Tal processo é conhecido como *habituação*. Se o estímulo é inicialmente muito desagradável (por ex., assustador ou nojento), ele tende a se tornar mais neutro ou a ser menos negativamente provocante; se o estímulo é inicialmente muito agradável, ele tende a se tornar mais neutro ou a ser menos positivamente estimulante. Nesse teste, 63 mulheres e 37 homens foram recrutados (100 universitários, no total).

Num experimento padrão de *habituação* (sem retrocausalidade), a sequência de eventos é a que se segue: 1º) o participante é repetidamente exposto de *forma subliminar* a uma imagem. Essa imagem é chamada *habituação-alvo*. 2º) ao participante é então mostrado duas imagens, lado a lado, e solicitado para indicar qual delas ele gosta mais. Uma das fotos é o *alvo*; a outra é uma imagem intimamente combinada que o participante não tinha visto antes. Se o participante prefere o *alvo*, a tentativa é marcada como um acerto. A taxa de acerto esperado pelo acaso é de 50%.

A versão com efeitos retroativos do protocolo acima simplesmente inverte as Etapas 1 e 2. Em cada tentativa, o participante é apresentado pela primeira vez a um par de fotos na tela do computador e é solicitado para indicar qual das imagens ele prefere. O computador, em seguida, seleciona aleatoriamente uma das duas fotos para servir como a *habituação-alvo* e a exibe *subliminarmente* várias vezes. O primeiro experimento de *habituação retroativa* foi composto de tentativas usando pares de imagens negativamente provocativas ou pares de imagens neutras como controle (pares de imagens positivamente estimulantes não foram introduzidos até o experimento 6). A hipótese de *habituação retroativa* era de que, sobre tentativas com pares de imagens negativas, os participantes iriam preferir o *alvo* do que o não-*alvo* mais do que 50% das tentativas.

A hipótese de *habituação retroativa* foi suportada. Nas tentativas com pares de imagens negativas, os participantes preferiram o *alvo* com frequência mais

Debate Psi
Uma Compilação da Evidência [Ano I, 2014]
www.debatepsi.com

significativa do que o não-alvo, 53,1 %, $t(99) = 2.23$, $p = 0.014$, tamanho do efeito (es) = 0.22.

- *A indução retroativa da sensação de tédio.* Uma afortunada descoberta foi observada em 10 sessões nas quais a taxa de acerto estava significativamente abaixo dos 50% atingidos pelas tentativas de controle, sugerindo que se poderia projetar uma experiência para induzir retroativamente nos participantes a sensação de tédio diante da exposição repetitiva a estímulos neutros (ou não provocantes).

Isto seria desejável pelo menos por duas razões. Em primeiro lugar, existem grandes diferenças etárias, culturais e individuais nas reações diante dos tipos de imagens negativas e eróticas usadas no experimento de *habituação retroativa*, tornando menos provável a replicação bem-sucedida através de deferentes populações. Em segundo lugar, alguns pesquisadores psi têm sido relutantes em realizar experimentos nos quais os participantes são expostos a esses tipos de imagens. O experimento agora apresentado testa diretamente a indução retroativa da sensação de tédio sobre estímulos neutros.

Num experimento de *mera exposição* que incluiu uma medida de "propensão ao tédio", Bornstein *et al.* (1990) relataram que apenas os participantes que não são tendentes ao tédio mostraram um *efeito de mera exposição* significativo (ou seja, aumentaram o seu gosto quando expostos muitas vezes ao estímulo). Isto sugere que o tédio dominou a habituação dos participantes propensos ao tédio. Bem (2011) informa que deveria ficar claro que os 2 itens diametralmente opostos na escala de pessoas que "*procuram estímulos*" ("Eu sou facilmente aborrecido" e "muitas vezes eu me divirto vendo filmes que eu já vi antes") também poderiam ser muito bem conceitualizados como um índice de propensão ao tédio.

Sendo assim, Bem, em primeiro lugar, testou a hipótese de que aqueles indivíduos que tiveram altos pontos na escala de "procurar estímulos" (ou seja, alta tendência ao tédio) mostrariam significativamente decréscimo 'do gostar' em relação ao alvo após repetidas exposições. 200 universitários participaram desse experimento, sendo 140 mulheres e 60 homens. O protocolo foi essencialmente o mesmo daquele utilizado nos experimentos de *habituação retroativa*. Em cada uma das 24 tentativas, foi exibido ao participante duas imagens combinadas, sendo-lhe solicitado escolher, entre elas, qual a imagem de sua preferência (o que fazia através de um clique de mouse). Depois da escolha, o computador selecionava aleatoriamente uma das duas imagens para servir como *alvo* e disparava flashes dela sobre a tela por 10 vezes. A *hipótese psi* era de que os participantes iriam atingir uma taxa de acerto significativamente mais baixa do que os 50% esperados pelo acaso (ou seja, iriam preferir com menor frequência a imagem-alvo, uma vez que os flashes subsequentes dela a tornariam, retrocausalmente, mais enfadonha aos olhos do participante com alta propensão ao tédio.

Ao contrário do que aconteceu nos experimentos de *habituação retroativa*, as exposições foram supraliminares (750-ms de duração seguidos de uma tela em branco por 250ms) e ampliadas de modo a preencher toda a tela. Daryl J. Bem imaginou que que as exposições supraliminares seriam mais susceptíveis de produzir a sensação de tédio depois de 10 exposições do que exposições subliminares utilizadas com sucesso nos experimentos originais de habituação retroativa. Além disso, porque 10 exposições

Debate Psi

Uma Compilação da Evidência [Ano I, 2014]

www.debatepsi.com

supraliminares em cada tentativa tomava bastante tempo, o número de tentativas foi reduzido a 24 para evitar que a experiência completa se tornasse enfadonha, e não apenas as imagens-alvo. A sequência aleatória dos pares de imagens, a localização das duas imagens (esquerda/direita) e a seleção do alvo foram implementadas randomicamente com o algoritmo PRNG Marsaglia.

Em todas as 200 sessões, a taxa de sucesso foi na direção correta, mas não significativamente diferente do acaso, 49,1 %, $t(199) = 1.31$, $p = 0.096$, *tamanho do efeito* (es) = 0.09. Por outro lado, pessoas que “procuram estímulos” foram positivamente correlacionadas com a performance psi (agora, através de baixas taxas de acerto), r (correlação) = 0.16, $p = 0.011$. Os participantes com alta “procura de estímulos” obtiveram uma taxa de acerto significativamente mais baixa do que o acaso, 47.9 %, $t(95) = 2.11$, $p = 0.019$, *tamanho do efeito* (es) = 0.22, binominal $z = -1.94$, $p = 0.026$, enquanto os participantes remanescentes, não, 50.1 %, $t(103) = 0.17$, $p = 0.43$.

Experimento com base no sistema de ‘pensamento lento’

A *facilitação retroativa da memória*. O presente experimento testou a hipótese de que a memória pode “funcionar em ambos os sentidos”, por meio de testes nos quais se exercitava um conjunto de palavras para torna-las mais fáceis de serem lembradas, ainda que o exercício (para decorar) acontecesse após o teste de recordação. Os participantes foram apresentados a um conjunto de palavras, sendo-lhes aplicado em seguida um teste (surpresa) de memória daquelas palavras. Após o teste, foram dados aos participantes alguns exercícios práticos sobre um subconjunto aleatoriamente selecionado daquelas palavras. A *hipótese psi* era de que os exercícios práticos iriam facilitar retroativamente a lembrança daquelas palavras e, conseqüentemente, os participantes iriam lembrar mais das palavras *a-serem-praticadas* do que em relação as palavras *não-praticadas*. 100 universitários participaram desse estudo, sendo 64 mulheres e 36 homens.

Aos participantes foram mostrados 48 substantivos, um de cada vez, durante 3 segundos cada. Essas palavras foram retiradas de quatro categorias (alimentos, animais, profissões e roupas) e foram apresentadas na mesma ordem para todos os participantes. Eles foram convidados a visualizar o referente de cada palavra à medida que elas apareciam na tela (por exemplo, se a palavra era *árvore*, eles visualizariam uma árvore). A eles então foi aplicado um teste (surpresa) de memória no qual foram solicitados a escrever todas as palavras que poderiam lembrar em qualquer ordem.

Após o participante completar o teste de memória, o computador selecionava aleatoriamente 6 palavras a partir de cada uma das 4 categorias para servir como palavras *a-serem-praticadas*, com as 24 palavras restantes servindo como palavras *não-praticadas* (controle). As 24 palavras *a serem-praticadas* apareciam todas juntas numa lista randomizada sobre a tela. O participante era informado de que as palavras foram extraídas de 4 categorias e solicitado a clicar sobre as 6 palavras de alimentos na lista (que se tingiam de vermelho quando clicadas) e, em seguida, digitar essas palavras em 6 espaços vazios na tela. A lista era reembaralhada e a mesma tarefa repetida para cada uma das outras três categorias de palavras. Ao todo, foi requerido ao participante para rastrear a lista de *palavras-a-praticar* por 4 vezes, clicando sobre as 6 palavras de cada categoria e escrevendo cada uma das 24 palavras.

Debate Psi
 Uma Compilação da Evidência [Ano I, 2014]
www.debatepsi.com

Diferente dos outros estudos, neste experimento cada palavra que o participante lembrava constituía uma tentativa e era pontuada ou como *palavra-praticada* ou *palavra-de-controle*. Por conseguinte, a mensuração de psi neste estudo é uma pontuação ponderada do diferencial de recordações (DR), definido como o número de *palavras-praticadas* lembradas menos o número de *palavras-controles* recordadas (P – C) multiplicado pela pontuação global de lembranças do participante (P + C). Isto dá mais peso aos participantes que recordaram mais palavras (“contribuíram com mais tentativas”) e é conceitualmente análogo à prática de estudos ponderados pelos *tamanhos da amostra* numa meta-análise. Para fins descritivos, a pontuação de DR é expressa como uma porcentagem da máxima pontuação possível de DR (576), que seria alcançado caso um participante recordasse todas as 24 *palavras-praticadas*, mas nenhuma das 24 *palavras-controle*. Assim, o %DR = [(P – C) x (P + C)] / 576, podendo variar de -100% a 100%, com pontuações positivas de %DR indicando que as *palavras-praticadas* foram mais recordadas do que as palavras de controle.

Os resultados mostram que a ação de exercitar um conjunto de palavras após o teste de memória, de fato, retroage no tempo para facilitar a recordação daquelas palavras: O %DR médio para a amostra total foi de 2.27%, $t(99) = 1.92$, $p = 0.029$ e *tamanho da amostra* (es) = 0.19. Mais uma vez, pessoas que "*procuram estímulos*" estavam significativamente correlacionadas com o desempenho psi (%DR): r (correlação) = 0.22, $p = 0.014$. Essa correlação é refletida nas fortes pontuações (%DR) de participantes com alta tendência a "*procurar estímulos*": 6.46%, $t(42) = 3.76$, $p = 0.0003$ e *tamanho da amostra* (es) = 0.57. Em contraste, aqueles com baixa propensão a "*procurar estímulos*" marcaram ao nível do acaso: -0.90%, $t(56) = -0.60$ e $p = 0.73$.

Resultados dos experimentos de ‘antecipação anômala de eventos futuros e aleatórios’

Na tabela abaixo, vemos as probabilidades e tamanhos do efeito totais para as amostras de cada um dos nove experimentos. Os testes envolveram mais de 1.000 participantes e todos os experimentos, exceto um (*indução retroativa da sensação de tédio*), atingiram resultados estatisticamente significativos. O tamanho do efeito médio da performance de psi, através dos nove estudos, foi de 0.22, com um Stouffer z de 6.66, $p = 2.68 \times 10^{-11}$ (Bem, Utts e Johnson, 2013).

FENÔMENO TESTADO E EXPERIMENTO	Probabilidade total da amostra	Tamanho do efeito total da amostra
<i>Aproximação e fuga precognitivas</i>		
<i>Deteção de estímulo erótico</i>	0.01	0.25
<i>Fuga de estímulo negativo</i>	0.009	0.20
<i>Priming retroativo</i>		
<i>Priming retroativo I</i>	0.007	0.26
<i>Priming retroativo II</i>	0.014	0.23
<i>Habituação retroativa</i>		
<i>Habituação retroativa I</i>		
<i>Tentativas negativas</i>	0.014	0.22
<i>Habituação retroativa II</i>		
<i>Tentativas negativas</i>	0.037	0.15
<i>Tentativas eróticas</i>	0.039	0.14
<i>Indução retroativa da sensação de tédio</i>	0.096	0.09
<i>Facilitação retroativa da memória</i>		

Debate Psi
Uma Compilação da Evidência [Ano I, 2014]
www.debatepsi.com

<i>Facilitação da memória I</i>	0.029	0.19
<i>Facilitação da memória II</i>	0.002	0.42
Resultado Global	2.68×10^{-11}	0.22

A meta-análise dos experimentos de ‘antecipação anômala de eventos futuros e aleatórios’

Em 2014, Daryl J. Bem, Patrizio Tressoldi, Thomas Rabeyron e Michael Duggan divulgaram uma meta-análise de 90 experimentos sobre antecipação anômala de eventos futuros e aleatórios [[acesse aqui](#)]. A meta-análise compreende todos os estudos do gênero achados entre 2000 a setembro de 2013, incluindo reportes não publicados, estudos par-revisados, artigos publicados em revistas do mainstream de psicologia, revistas especializadas e atas de conferências. Estudos também foram buscados no Google Scholar, PubMed e PsycInfo, além de terem sido contatados pesquisadores psi através de e-mail e da academia.edu.

93 estudos foram encontrados nos quais os autores da meta-análise poderiam assegurar o número de participantes e o *tamanho do efeito*. Dois foram eliminados porque o tamanho da amostra era irrisória (1 e 9 participantes). O terceiro, com resultado positivo, possuía muitas análises *post-hoc*, sendo mais exploratório do que confirmatório do fenômeno analisado. Assim, o banco de dados final ficou com 90 experimentos.

Patrizio Tressoldi e Thomas Rabeyron, de forma independente, categorizaram cada estudo com as seguintes variáveis: a) tipo(s) de efeito(s) testado(s); b) número de participantes envolvidos no estudo; c) estatística descritiva ou inferencial necessária para calcular o *tamanho do efeito*; d) se o experimento tinha sido par-revisado; e) o tipo de replicação.

Nessa última variável (e), cada experimento foi separado dentro de 3 categorias: I) uma exata replicação do experimento de Bem; II) uma replicação modificada; ou III) um projeto experimental independente que avaliava, de uma maneira alternativa, a habilidade de antecipar eventos futuros selecionados randomicamente. 69 experimentos foram classificados nas categorias I e II, enquanto 11 foram classificados na categoria III.

O banco de dados completo foi composto por 90 experimentos conduzidos entre 2002 e 2013, originados de 33 laboratórios diferentes localizados em 14 países e envolveram 12.406 participantes. Os experimentos originais de Bem foram incluídos nesta meta-análise de modo atingir os 90 estudos analisados. 51 (56,6%) das experiências tinham sido publicadas em revistas par-revisadas ou em atas de conferências. O banco de dados completo dos experimentos com os correspondentes *tamanhos de efeito* e as atribuições de categoria é apresentado abaixo, com um gráfico sobre o *tamanho de efeito* individual e seus 95% de intervalo de confiança (IC). “*A principal questão dirigida para esta meta-análise é se o banco de dados fornece evidência global para a existência da antecipação anômala de eventos futuros e aleatórios(...). A resposta é sim: O tamanho de efeito global (g Hedges) é 0.09, z combinado = 6.33, p = 1,2 × 10⁻¹⁰. O valor de Bayesiano BF é 1.24 × 10⁹, absurdamente excedendo assim o valor critério de 100 o qual*

Debate Psi
 Uma Compilação da Evidência [Ano I, 2014]
www.debatepsi.com

é considerado como constituindo 'prova decisiva' para a hipótese experimental (Jeffreys, 1961)".

Uma questão subsidiária é se investigadores independentes podem replicar com êxito os experimentos originais de Bem (2011). A resposta novamente é sim: quando as experiências de Bem são excluídas, o *tamanho do efeito* para as replicações é 0.07, z combinado = 4.25, $p = 1,1 \times 10^{-5}$ e o valor BF é 757, que, de novo, excede absurdamente o valor critério de 100 para 'prova decisiva' (esta análise também exclui as 11 experiências que não foram projetadas para replicar as experiências de Bem). Quando as próprias experiências de Bem são excluídas, as replicações exatas e as replicações modificadas produzem resultados estatísticos comparáveis, com as replicações exatas produzindo *tamanhos de efeito* um pouco maiores (replicações exatas: *tamanho do efeito* (g Hedges) = 0.09 e $p = 1.39 \times 10^{-3}$. Replicações modificadas: *tamanho do efeito* (g Hedges) = 0.05 e $p = 1.31 \times 10^{-3}$).

<i>Resultados meta-analíticos para todos os experimentos e para as replicações independentes dos experimentos de Bem</i>						
	Número de experimentos	Número de participantes	Tamanho do efeito (g Hedges)	IC 95%	z combinado ou Fator Bayesiano (BF _{h1/h0})	p
Todos os experimentos	90	12.406	0.09	[0.06, 0.12]	$z = 6.33$	1.2×10^{-10}
Análise Bayesiana			0.09	[0.03, 0.15]	BF = 1.24×10^9	
Replicações independentes dos experimentos de Bem	69	10.082	0.07	[0.04, 0.10]	$z = 4.45$	1.07×10^{-5}
Análise Bayesiana			0.07	[0.01, 0.15]	BF = 1,907	
Replicações exatas	31	2.082	0.09	[0.03, 0.14]	$z = 2.99$	1.39×10^{-3}
Replicações modificadas	38	8.000	0.05	[0.02, 0.09]	$z = 3.01$	1.31×10^{-3}

Na tabela abaixo, vemos que experimentos de '*pensamento rápido*' claramente revelam melhores resultados do que as experiências de '*pensamento lento*': cada protocolo de '*pensamento rápido*' alcançou, individualmente, um efeito estatisticamente significativo, com um tamanho de efeito global de 0.11 e um z combinado maior que *sigma* 7. Em contraste, os experimentos de '*pensamento lento*' alcançaram um tamanho de efeito global de apenas 0.03, falhando em atingir um nível convencional de significância estatística ($p = .20$).

<i>Resultados meta-analíticos em função do protocolo e tipo de experimento</i>						
Tipo de experimento	Número de experimentos	Número de participantes	Tamanho do efeito	IC 95%	z combinado	p
<i>Protocolo de 'pensamento rápido'</i>						
1. Detecção precognitiva de reforço (estímulo erótico)	14	863	0.14	[0.08, 0.21]	4.22	1.2×10^{-5}
2. Fuga precognitiva de estímulo negativo	8	3,120	0.09	[0.03, 0.14]	3.10	.002
3. Priming retroativo	15	1,154	0.12	[0.04, 0.20]	2.94	.003
4. Habituação retroativa	20	1,780	0.11	[0.04, 0.17]	3.32	.001
5. Prática retroativa	4	780	0.10	[0.04, 0.18]	3.02	.003
Todos os experimentos de 'pensamento rápido'	61	7,697	0.11	[0.08, 0.14]	7.15	4.3×10^{-13}
<i>Protocolo de 'pensamento lento'</i>						

Debate Psi
 Uma Compilação da Evidência [Ano I, 2014]
www.debatepsi.com

1. Facilitação retroativa da prática sobre a memória	27	4,601	0.04	[-0.01, 0.09]	1.66	.10
2. Facilitação retroativa da prática sobre o texto lendo rápido	2	108	-0.14	[-0.54, 0.25]	-0.72	.47
Todos os experimentos de 'pensamento lento'	29	4,709	0.03	[-0.02, 0.08]	1.25	.20

Os autores da meta-análise levantam duas razões para que os experimentos de '*pensamento lento*' não tenham atingido a significância estatística. Primeiro, das 27 tentativas de replicações do estudo de *facilitação retroativa da memória*, 11 não aderiram ao protocolo original de Bem. Quando consideradas somente as 16 replicações exatas desse protocolo, elas atingem um *tamanho do efeito* global comparável àquele dos experimentos de '*pensamento rápido*'. Segundo, os autores acreditam que o protocolo de experimentos de '*rápido pensamento*' é mais suscetível de produzir evidência para psi, porque previne a ação de estratégias conscientes e cognitivas as quais interferem com a natureza implícita, inconsciente e automática do funcionamento psi.

O efeito *filedrawer* é absurdamente inverossímil para explicar os resultados altamente positivos em favor de psi. Seriam necessários 1.513 estudos não publicados e com resultados negativos para reduzir a significância estatística a níveis do acaso. Ou seja, seriam necessários 17 estudos não publicados para cada estudo incluído na meta-análise, o que excede e muito a proporção científica e razoavelmente aceita de 5:1. Além do mais, conforme já informado, os autores efetivaram uma minuciosa busca por estudos do tipo *presentimento*.

Uma outra objeção que se pode fazer seria a de que na meta-análise foram incluídos experimentos não submetidos ao filtro da par-revisão, logo, eventuais falhas metodológicas desses estudos podem ter passado despercebidas pelos respectivos autores e talvez provocassem uma inflação estatística favorável a psi. Olhando mais de perto, essa contestação não tem muita força. Vejamos. Não houve uma diferença estatística significativa entre os grupos de estudos par-revisados (51) e não par-revisados (39). Aliás, os estudos par-revisados produziram resultados ligeiramente mais favoráveis a psi do que em relação aos que não foram submetidos ao processo de revisão por pares (1.6×10^{-7} contra 1.5×10^{-6}). Além do mais, a própria utilização de estudos que não foram submetidos ao processo de par-revisão fortalece o resultado favorável a psi contra a possibilidade do *filedrawer*.