

# Conhecimento comum, coordenação e a lógica das emoções autoconscientes

Kyle A. Thomas,<sup>a,\*</sup> Peter DeScioli,<sup>b</sup> Steven Pinker<sup>a</sup>

<sup>a</sup> - Departamento de Psicologia, Universidade de Harvard, EUA

<sup>b</sup> - Departamento de Ciências Políticas, Universidade de Stony Brook, EUA

Informações sobre o Artigo:

Recebido originalmente em 22 de novembro de 2016

Revisão final recebida em 3 de dezembro de 2017

## 1. Introdução

Imagine-se derrubando um prato de comida no colo diante uma multidão. Depois, talvez você fixe seu olhar no seu celular para evitar admitir a trapalhada diante dos espectadores. Da mesma forma, depois de decepcionar sua família ou seus colegas, pode ser difícil olhá-los nos olhos. Por que as pessoas evitam admitir suas gafes ou transgressões que elas sabem que a audiência já conhece?

Depois de uma transgressão, as pessoas sentem as emoções autoconscientes negativas de vergonha, embaraço ou culpa, e essas emoções as ajudam a regular suas relações (Beer, Heerey, Keltner, Scabini, & Knight, 2003; Keltner & Buswell, 1997; Ketelaar & Au, 2003; Miller, 1995; Steckler & Tracy, 2014; Sznycer et al., 2016; Tangney & Tracy, 2012; Tracy & Robins, 2004). Um transgressor demonstrou inépcia, o que pode prejudicar sua reputação como cooperador valioso ou pode demonstrar desprezo pelo bem-estar de alguém, o que pode prejudicar sua reputação como cooperador digno de confiança. O desconforto causado pelas emoções resultantes, mesmo quando sentidas privadamente, motiva a pessoa a gerenciar essas ameaças, atraindo sua atenção para a transgressão e motivando-a a fazer correções e evitar atos semelhantes no futuro (Ketelaar & Au, 2003; McCullough, 2008; Sznycer et al., 2016; Trivers, 1971).<sup>1</sup>

A ideia que as emoções autoconscientes regulam as relações também explica por que a presença de uma audiência intensifica as sensações de embaraço, vergonha e culpa (Bosch et al., 2009; Gruenewald, Kemeny, Aziz, & Fahey, 2004; McCullough, 2008; Modigliani, 1971; Smith, Webster, Parrot, & Eyre, 2002; Tangney, Miller, Flicker, & Barlow, 1996; Wolf, Cohen, Panter, & Insko, 2010). Se os espectadores inferem que uma transgressão é resultado de uma disposição estável que indica futura incompetência ou abuso, agora eles têm motivos para desvalorizar, condenar ao ostracismo ou punir o transgressor. Para evitar essas consequências prejudiciais, o transgressor precisa persuadir os espectadores que o ato não foi intencional e, portanto, não é representativo de sua disposição subjacente, ou que ele mudará sua disposição e não repetirá o comportamento no futuro. Além disso, para que tais garantias sejam mais que apenas conversa fiada e egoísta, elas precisam soar confiáveis: O transgressor deve sofrer um custo na forma de desconforto visível e talvez reparação concreta, bem como exibir sinais que a alteração nas prioridades é produto de emoções involuntárias em vez de cálculos estratégicos conscientes. Na realidade, a pesquisa sobre psicologia da contrição e

---

<sup>1</sup> Uma análise de transgressões sociais requer diferenciação precisa das partes individuais em interações diáticas. O uso de *ele* ou *ela* e *eles* genericamente em tais discussões pode rapidamente se tornar desajeitado ou ambíguo (Pinker, 2014). Para evitar confusão, iremos nos referir a um ator hipotético usando pronome masculino singular e a espectadores hipotéticos usando pronomes no plural (Pinker, 2014).

perdão mostra que as emoções negativas autoconscientes têm essas especificações (Dijk, de Jong, & Peters, 2009; Dijk, Koenig, Ketelaar, & de Jong, 2011; Frank, 1988; Keltner & Buswell, 1996; Ketelaar & Au, 2003; McCullough, 2008; McCullough, Kurzban, & Tabak, 2010; Tracy & Robins, 2004; Trivers, 1971).

Como notamos no início, a intuição sugere que as emoções autoconscientes têm mais uma característica: elas são sentidas ainda com maior intensidade quando o ator reconhece perante a audiência que ele está consciente que a audiência tem conhecimento de sua transgressão (por exemplo, olhando nos olhos deles). Neste artigo, tentamos tornar precisa esta hipótese, testar se é correta e explicá-la.

### *1.1. A teoria de jogo e a psicologia do conhecimento comum*

A diferença entre evitar e reconhecer eventos pouco lisonjeiros se assemelha à distinção em teoria de jogo entre *conhecimento compartilhado* e *comum*. O conhecimento comum é uma recorrência infinita de estados de conhecimento social, tais como A sabe x, B sabe x, A sabe que B sabe x, B sabe que A sabe x, A sabe que B sabe que A sabe x, A sabe que B sabe que A sabe que B sabe x, e assim indefinidamente. Em contraste, o conhecimento compartilhado se refere a qualquer nível menor, finito de estados de conhecimento social (por ex., A sabe que B sabe que A conhece x, e nada mais).

O conhecimento comum tem um papel importante em *jogos de coordenação* — interações com múltiplos equilíbrios em que a melhor jogada de cada jogador depende da jogada do outro (Lewis, 1969; Schelling, 1960; Skyrms, 2004). O caso paradigmático é o jogo *Stag Hunt* (Caça a Veados), em que dois jogadores podem caçar coelhos individualmente (com prêmio baixo) ou podem trabalhar juntos para caçar um veado (prêmio alto). Porém, se um jogador caçar um veado sozinho a caça falha e ele não ganha nada. Para se coordenarem, cada jogador precisa saber que o outro jogador sabe que há uma oportunidade de caçar um veado, e mais, que essa informação é de conhecimento comum. Formas menores de conhecimento compartilhado, em que cada jogador sabe que os veados estão correndo, mas se pergunta se o outro jogador sabe disso ou se sabe que ele sabe, podem levar facilmente a uma custosa falta de coordenação (Chwe, 2001; Dalkiran, Hoffman, Paturi, Ricketts, & Vattani, 2012; Rubinstein, 1989; Schelling, 1960).

Pesquisas psicológicas recentes demonstram que as pessoas realmente entendem o conhecimento comum como qualitativamente distinto do conhecimento compartilhado e as decisões dessas pessoas diferem com base nessa distinção (Thomas, DeScioli, Haque, & Pinker, 2014). Thomas et al. (2014) colocou participantes em cenários hipotéticos com os prêmios de um *Stag Hunt*: eles tinham que decidir se preferiam trabalhar sozinhos, com prêmio garantido mas inferior, ou se tentariam trabalhar junto com um parceiro para obter prêmio mais alto que eles só receberiam se o parceiro fizesse a mesma escolha. A fonte da informação sobre os prêmios variava, às vezes permitindo apenas conhecimento privado, às vezes conhecimento compartilhado (transmitido por um menino de recados não confiável), às vezes conhecimento comum (transmitido por um alto-falante). Os participantes tinham mais probabilidade de preferir entrar em coordenação, conseguindo assim um prêmio mais alto, quando tinham conhecimento comum das oportunidades, em linha com as previsões teóricas de jogo. Logicamente, ninguém pode conceber explicitamente uma proposição infinitamente integrada; presumivelmente, o conhecimento comum é concebido nas mentes das pessoas como uma fórmula mental finita ou símbolo, correspondendo à intuição que algo é “público” ou “está ali perto”.

O conceito que as pessoas têm do conhecimento comum afeta não somente o raciocínio estratégico delas, mas também uma variedade de outros fenômenos psicológicos incluindo discurso indireto (Chakroff, Thomas, Haque, & Young, 2015; Lee & Pinker, 2010), julgamento moral (DeScioli & Kurzban, 2013), ignorância pluralística (Willer, Kuwabara, & Macy, 2009), e difusão de responsabilidade (Thomas, De Freitas, DeScioli, & Pinker, 2016). Investigamos o papel do conhecimento comum nas emoções autoconscientes. Especificamente, propomos que, em comparação com o conhecimento compartilhado, o conhecimento comum de uma transgressão impõe um imperativo maior de retificá-la, o que, por sua vez, causa emoções autoconscientes sentidas mais intensamente.

### *1.2. Transgressões sociais, sinais sociais e conhecimento compartilhado*

Começamos com a ideia que uma ofensa representa uma ameaça maior à reputação do transgressor se uma audiência não somente sabe privadamente sobre a transgressão, mas também sabe que o transgressor sabe que eles sabem. Alguém que fracassa em uma tarefa ou decepciona um amigo ficará mal visto se não pedir desculpa, mas vai ficar ainda pior se a audiência souber que ele sabe que a audiência presenciou a ofensa e ainda não há pedido de desculpa à vista.

O raciocínio é o seguinte. Um transgressor que deixa de se desculpar quando uma audiência observa a ofensa, sem conhecimento compartilhado, com certeza já fica em uma posição comprometida. Mas a audiência ainda poderia estar incerta sobre as intenções do transgressor. Dessa forma, do ponto de vista da audiência, ainda é possível que o transgressor pretenda comportar-se de forma diferente no futuro. Essa incerteza dá aos espectadores razão para não responder draconianamente (já que eles também podem perder os benefícios de futura cooperação com o ator), e essa pressão moderada sobre o transgressor deve provocar formas brandas de embaraço, culpa e vergonha.

Em contraste, um transgressor que não pede desculpas, nem mesmo quando os espectadores sabem que ele sabe que eles sabem que a ofensa transmite informações adicionais: que ele não está disposto ou que não consegue mudar seu comportamento para preservar sua posição diante dos espectadores. Em outras palavras, ele estaria sinalizando que está preparado para abandonar totalmente a cooperação e renunciar a todos os benefícios dessa cooperação: a estratégia de um psicopata, lobo solitário, um excêntrico ou maluco (consultar também Goffman, 1959, 1978). Para um membro de uma espécie altamente social, as consequências de tal exibição podem ser graves e assim sugerimos que as emoções autoconscientes se adaptam para evitar tais consequências.

Além disso, as pesquisas demonstraram que os transgressores tendem a pedir desculpas e confessar estrategicamente (Sznycer, Schniter, Tooby, & Cosmides, 2015). Especificamente, os transgressores tendem a confessar e pedir desculpas quando fazer isso os beneficia mais, que é quando: (1) o custo da transgressão é baixo para a vítima; (2) o benefício da transgressão é alto para eles; e (3) as vítimas já podem saber da transgressão ou podem facilmente descobrir sobre ela de alguma outra forma. Assim, as vítimas enfrentam um problema de detecção de sinal quando recebem um pedido de desculpas: É um sinal honesto ou é lero-lero estratégico? Principalmente, quando a vítima acha que o transgressor acha que está revelando a transgressão à vítima pela primeira vez, em vez da vítima já saber o que o transgressor fez, então o pedido de desculpas do transgressor parece mais crível, já que revela a ofensa em vez de só remendar as coisas depois que a ofensa já foi descoberta (Sznycer et al., 2015; Utikal, 2012; Weiner, Graham, Peter, & Zmuidinas, 1991). Assim, logo que uma transgressão se torna conhecimento compartilhado, ambas as partes poderiam, razoavelmente, supor que a expectativa da vítima de receber um pedido de desculpas aumenta, assim também a janela de tempo do transgressor para fornecer uma desculpa se estreita da mesma forma. Em

tais casos, pode ser necessário mais do que apresentar uma desculpa para torná-la crível, e os sinais emocionais incontroláveis produzidos por emoções autoconscientes podem aumentar a credibilidade do que, de outra forma, poderia ser percebido como simples lero-lero.<sup>2</sup>

Esta teoria de gerenciamento de reputações de emoções negativas autoconscientes é consistente com o fato que essas emoções são evocadas por dois tipos diferentes de transgressão: aquelas que mostram desprezo pelo bem-estar de alguém e as que revelam incompetência. Ambas as categorias decorrem de teorias de cooperação e escolha de parceiros, que demonstram como os indivíduos ganham mais com a cooperação quando preferem parceiros que valorizem seu bem-estar e que sejam suficientemente competentes para retribuir (Tooby & Cosmides, 1996; Trivers, 1971). Os transgressores que não fazem reparações têm mais probabilidade de tornar a ofender no futuro, seja devido a indiferença, incompetência, hostilidade ou ignorância. Sugerimos que a falta de contrição é ainda mais reveladora quando a audiência sabe que o ator sabe que a audiência testemunhou a transgressão.

### *1.3. Relações, coordenação e conhecimento comum*

Além disso, propomos que o conhecimento comum de uma transgressão é ainda pior para o ator do que o conhecimento compartilhado. O conhecimento compartilhado de uma ofensa ameaça prejudicar a reputação de um ator; o conhecimento comum acrescenta a ameaça adicional de desestabilizar uma relação valiosa.

De uma perspectiva de teoria de jogos, as relações podem ser vistas como jogos de coordenação (Dalkiran et al., 2012; Lee & Pinker, 2010; Pinker, Nowak, & Lee, 2008), já que com frequência é mutuamente vantajoso aparentar bondade, respeito, apoio, indiferença, afeição ou maldade para com os que demonstram a mesma coisa para nós. Ou seja, parceiros sociais geralmente se beneficiam com a coordenação sobre o mesmo tipo de relação, e as incompatibilidades de relações podem custar caro. As pessoas podem usar sinais presentes e históricos passados para estabelecer relações cooperativas. Mesmo assim, cada parte pode abandonar unilateralmente uma relação a qualquer momento. Assim, cada parte se beneficia com a checagem periódica do estado da relação e com a reafirmação de sua compreensão da relação para sua contraparte.

Como as relações dependem de crenças e expectativas sincronizadas, as relações são frágeis, podendo ser minadas por transgressões intencionais ou até acidentais. Uma amizade pode esfriar e desconhecidos podem migrar de neutros para antagônicos. Nesses casos, os indivíduos ficam na difícil situação de adivinhar os estados mentais dos demais para prever o comportamento futuro deles, e uma transgressão não admitida pode destruir uma relação direcionando as expectativas de uma parte em direção a não-cooperação no futuro.

Esse problema é bem estudado no contexto de conflito violento em que rivais podem acidentalmente entrar em uma “armadilha Hobbesiana”, como uma corrida armamentista desenfreada (Pinker, 2011; Schelling, 1960). Por exemplo, duas nações sem vínculos próximos poderiam preferir evitar uma guerra depois de uma disputa, mas cada uma delas pode mobilizar forças militares caso a outra decida atacar, e isso, por sua vez, incentiva cada uma a atacar preventivamente para evitar perder a vantagem de atacar primeiro.

Essa mesma dinâmica pode ocorrer em relações interpessoais. Um amigo, cônjuge, colega ou desconhecido pode ignorar uma pessoa simplesmente por temer o mesmo da outra pessoa. Quando ocorre uma transgressão pública — roubo de dinheiro, suspeita de infidelidade, uma calúnia ou um prato de comida derrubado — os indivíduos podem tentar adivinhar suas

---

<sup>2</sup> Agradecemos um revisor anônimo por este insight.

relações e acabar em um tipo de “armadilha de deserção” na qual cada um deles enxerga a relação mais negativamente apenas porque pensam que a outra pessoa fez o mesmo.

Para fugir dessas armadilhas, os parceiros precisam de algum modo reafirmar a relação garantindo entre si que ambos pretendem mantê-la. Propomos que as emoções autoconscientes destinam-se a facilitar essas reafirmações: ajudam um transgressor a sinalizar contrição de forma imediata e pública a fim de impedir uma espiral de deterioração do relacionamento. Consistentes com essa ideia, pesquisas têm mostrado que o embaraço e o rubor das faces podem servir como sinais de contrição e ajudam a apaziguar uma audiência potencialmente hostil (Dijk et al., 2009; Dijk et al., 2011; Keltner, 1995; Keltner & Buswell, 1997), e todas as três emoções autoconscientes negativas motivam as pessoas a tomarem medidas corretivas e fazer correções (Beer et al., 2003; Keltner & Buswell, 1997; Ketelaar & Au, 2003; Tracy & Robins, 2004).

Essencialmente, a lógica da coordenação implica que o conhecimento comum de uma transgressão pode ser mais prejudicial para uma relação do que até mesmo níveis arbitrariamente altos de conhecimento compartilhado. Alinhando as crenças e expectativas dos indivíduos, o conhecimento comum de uma transgressão pode forçar os parceiros sociais a reconhecerem verdades desagradáveis e criarem uma armadilha de deserção que pode destruir sua relação caso não seja abordada. Em contraste, se uma transgressão social for meramente conhecimento compartilhado, os parceiros sociais não são necessariamente forçados a reconhecê-la, o que pode mantê-los afastados de uma armadilha de deserção. Uma pequena quantia de incerteza significa que deixar de pedir desculpas não é necessariamente uma afronta direta que desafia o status quo, mas, em vez disso, poderia ser uma tentativa de evadir-se da ofensa. Embora a vítima ainda possa preferir desculpas à evasão, esta última pelo menos sugere que o ator pretende manter o status quo. Em resumo, o conhecimento comum de uma transgressão força as pessoas a revisitarem, renegociarem, e/ou reafirmarem os termos de sua relação; ao passo que, até níveis elevados de conhecimento compartilhado proporcionam espaço de manobra que permite as pessoas simplesmente ignorarem a infração se assim desejarem (Dalkiran et al., 2012 fornecem um modelo formal desse fenômeno).

Finalmente, as ameaças colocadas pelo conhecimento comum também podem ir além da relação em si e incluir terceiros que observam ou descobrem a infração. Quando uma infração é conhecimento comum, os terceiros têm maior capacidade para condenar e punir o infrator (DeScioli & Kurzban, 2013). Portanto, o conhecimento comum representa uma ameaça especial vinda de terceiros, além do dano que pode fazer dentro dos relacionamentos. Aqui focamos em relações diáticas, mas voltaremos à questão da coordenação de terceiros na Discussão Geral (Seção 4.3).

Em resumo: uma relação cooperativa é ameaçada por uma transgressão que é de conhecimento comum (a menos que seja especificamente corrigida), mas é menos ameaçada por uma transgressão que seja meramente conhecimento compartilhado (que não precisa necessariamente ser corrigida).

#### *1.4. Como as emoções autoconscientes ajudam os indivíduos a gerenciarem relações?*

As principais teorias das emoções autoconscientes sustentam que elas funcionam para ajudar os indivíduos a gerenciarem relações e navegarem em seu ambiente social (ver Beer et al., 2003; Dijk et al., 2009; Keltner, 1995; Keltner & Buswell, 1997; Ketelaar & Au, 2003; McCullough, 2008; Miller, 1995; Steckler & Tracy, 2014; Sznycer et al., 2016; Tangney & Tracy, 2012; Tracy & Robins, 2004). Especificamente, essas teorias focam na maneira como as emoções autoconscientes motivam as pessoas a evitarem transgressões, evitam que informações prejudiciais se espalhem e acalmam as outras pessoas depois de ocorrer uma transgressão. A descoberta que as emoções autoconscientes são sentidas com maior

intensidade quando há uma audiência presente proporciona suporte empírico para este conjunto de ideias.

No entanto, embora manter a reputação de uma pessoa seja um aspecto importante do gerenciamento de relações, nossa análise sugere outro elemento crítico que as teorias existentes não abordaram: o estabelecimento e manutenção da coordenação. Se relações exigem coordenação, então as emoções autoconscientes serão sentidas com maior intensidade quando uma transgressão for reconhecida frente a uma audiência, em comparação com quando ela é deixada não reconhecida, porque o conhecimento comum da ofensa coloca uma ameaça à coordenação.

Assim, a teoria da coordenação das emoções autoconscientes estima que o conhecimento comum de uma transgressão vai desencadear sensações mais intensas de embaraço, vergonha ou culpa do que se a transgressão for meramente conhecimento compartilhado entre uma audiência e o transgressor. As teorias de reputação propostas em pesquisas anteriores não prevêem claramente esse outro efeito do conhecimento comum. A reputação de um transgressor pode ser prejudicada logo que alguma outra pessoa descobrir sobre a transgressão dele. Dessa forma, saber que uma audiência está consciente da transgressão é suficiente para saber que a reputação da pessoa está em perigo. O conhecimento comum não precisa aumentar esse perigo porque, uma vez que um transgressor sabe que um terceiro está consciente da ofensa e, mesmo assim, ainda não a aborda, o prejuízo à reputação já foi feito. Propomos que, como as emoções autoconscientes rastreiam preocupações sobre a reputação e a coordenação, elas deveriam ser sentidas mais intensamente com conhecimento compartilhado do que com conhecimento privado, e ainda com maior intensidade com conhecimento comum do que com conhecimento compartilhado.

### *1.5. Testes empíricos*

Testamos a teoria da coordenação em dois experimentos, um usando cenários hipotéticos e outro usando uma situação real, embaraçosa, em que os participantes cantam para uma audiência de desconhecidos. Um problema para testar qualquer teoria de emoções autoconscientes negativas de embaraço, culpa e vergonha é como distingui-las de outras emoções com valência negativa tal como raiva, tristeza ou ansiedade. Uma das descobertas mais antigas da ciência afetiva é que as diferentes emoções estão posicionadas ao longo de um pequeno número de dimensões, sendo que as positivas-negativas estão entre as mais destacadas (Rubin & Talarico, 2009; Russell & Barrett, 1999; Watson, Clark, & Tellegen, 1988). Além disso, as diferentes emoções no pólo negativo se encaixam em uma classe natural, podendo compartilhar fenomenologia inerente, substratos fisiológicos e patologias (como na comorbidade de depressão, ansiedade e fobias sociais) (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001). Assim, ao testar a hipótese que busca identificar os gatilhos cognitivos e sequelas comportamentais de emoções específicas, é essencial distinguir a experiência e relato dessas emoções a partir de uma negatividade mais global que pode ser desencadeada por uma variedade de situações estressantes não específicas. Nos dois experimentos, tentamos assim distinguir emoções autoconscientes de negatividade mais geral subtraindo um compósito de emoções autoconscientes reportadas pelos participantes de suas emoções autoconscientes reportadas.

No primeiro experimento, os participantes leram frases na primeira pessoa descrevendo situações com uma transgressão que tipicamente gera emoções autoconscientes negativas, raiva ou tristeza. Escolhemos raiva e tristeza como casos de emoções básicas bem estudadas (Ekman, 1999; Tracy & Robins, 2004). Em diferentes versões dos cenários, variamos os níveis de conhecimento das personagens sobre a transgressão como conhecimento privado,

compartilhado ou comum. Foi pedido aos participantes que se imaginassem no cenário e avaliassem quanto eles sentiriam emoções diferentes.

Essencialmente, notamos que as avaliações dos participantes para emoções em cenários hipotéticos podiam refletir uma mistura das emoções reais de outras pessoas que os participantes podem estar experimentando indiretamente, bem como suas crenças sobre quais emoções elas prevêm que sentiriam em tais cenários (que podem ou não ser exatas). Cenários escritos podem potencialmente evocar emoções genuínas. Por exemplo, as pessoas comumente relatam retração, medo e outras reações emocionais ao serem solidários com uma personagem fictícia em uma situação estressante. Essas emoções indiretas e imaginadas são o alicerce para teorias baseadas em imaginário tais como dessensibilização e impropriedade (Holmes, Arntz, & Smucker, 2007; Mar & Oatley, 2008). No entanto, não podemos distinguir emoções genuínas de emoções previstas em auto-relatos, embora esperemos que elas estejam altamente correlacionadas.

Se, conforme previsto pela teoria da coordenação, reconhecer uma transgressão é um gatilho particularmente potente para emoções autoconscientes, então os participantes reportarão emoções autoconscientes mais intensas quando as transgressões forem de conhecimento comum em comparação a conhecimento compartilhado. Como se explicou acima, a teoria da coordenação vai mais além de apenas prever que as emoções autoconscientes devem ser sensíveis a esses níveis variantes de conhecimento; ela também prevê que esses efeitos serão relativamente específicos para as emoções autoconscientes, em oposição a outras emoções negativas que têm importantes funções adicionais além de qualquer papel que possam ter no gerenciamento de relações. Portanto, a teoria da coordenação prevê que os níveis de conhecimento: (a) afetam as avaliações de emoções autoconscientes depois de subtrair o efeito negativo geral; e, (b) afetam as emoções autoconscientes mais do que as emoções básicas tradicionais de raiva e tristeza.

O segundo experimento vai ainda mais além, colocando realmente os participantes em uma situação embaraçosa: karaokê. Os participantes cantaram em um Karaokê para um júri em uma sala separada conectada por transmissão de vídeo bidirecional e lhes foi dito que o júri sabia que o participante sabia que o júri estava assistindo pela câmera (conhecimento comum), ou lhes disseram que o júri achava que estava assistindo secretamente a apresentação, sem o conhecimento do participante (conhecimento compartilhado). Os participantes então reportaram com que intensidade sentiram 21 emoções diferentes durante sua apresentação. A teoria da coordenação prevê maior embaraço com conhecimento comum do que com conhecimento compartilhado e, mais especificamente, que os sentimentos distintos de embaraço dos participantes depois de subtrair o efeito negativo geral, serão maiores com conhecimento comum do que com conhecimento compartilhado.

## **2. Experimento 1**

Neste experimento os participantes relataram como reagiriam a diferentes níveis de conhecimento em um de cinco cenários. Três dos cenários foram projetados para evocar emoções autoconscientes negativas: ser apanhado zombando de um amigo, flatulência audível durante uma palestra, ou reportar compras desonestamente para conseguir reembolso. Os outros dois cenários foram projetados para evocar emoções negativas não autoconscientes: tristeza ao descobrir que o companheiro romântico da pessoa está prestes a terminar o relacionamento e raiva por descobrir que um conhecido estava tentando sabotar uma amizade valiosa. Para cada cenário, os participantes avaliavam quanto sentiriam de uma bateria de emoções. As avaliações das múltiplas emoções nos permitiram examinar as emoções

autoconscientes depois de subtrair o efeito negativo geral, distinguindo as primeiras das últimas.

Cada participante lia cinco frases diferentes (apresentadas em ordem aleatória) que descreviam diferentes níveis de conhecimento do evento focal em um dos cenários. Ou seja, o cenário variava entre assuntos; ao passo que o conhecimento variava dentro dos assuntos, indo de *conhecimento privado*, em que só o participante tinha conhecimento do evento focal, até *conhecimento compartilhado*, em que o participante e outra pessoa sabiam sobre o evento focal, mas nenhum deles estava totalmente informado do conhecimento da outra pessoa, para *conhecimento comum* em que o evento focal era conhecimento sabido em comum entre o participante e outra pessoa.

Como as emoções autoconscientes negativas são tipicamente causadas pelas próprias ações da pessoa, ao passo que outras emoções negativas tendem a ser evocadas pelas ações de outras pessoas, os eventos focais nos cenários correspondentes precisavam ser diferentes: os participantes assumiam o papel do ator nos cenários autoconscientes e de um observador nos cenários básicos de emoções. Como [Tangney e Tracy \(2012\)](#) explicam essa distinção, as emoções autoconscientes diferem de outros tipos de emoções porque "... [as emoções autoconscientes] fundamentalmente envolvem as reações das pessoas a suas próprias características ou comportamento" (pg. 446). Em contraste, é difícil pensar em fazer algo, nós mesmos, que evoque *diretamente* emoções não autoconscientes tais como raiva ou tristeza (em oposição a iniciar uma cadeia de eventos imprevisíveis, o último dos quais pode ser a causa imediata da emoção).

## 2.1. Método

### 2.1.1. Participantes

Conforme uma análise de poder para detectar efeitos médios, cem participantes dos EUA foram recrutados do Mechanical Turk da Amazon para cada um dos cinco cenários ( $N = 500$ , na faixa etária de 18 a 81 anos), recebendo de 2 a 3 dólares por sua participação. Como a atenção é uma preocupação importante com participantes remotos, incluíram-se checagens de compreensão e atenção; qualquer participante que falhasse em qualquer um desses itens era excluído (ver Procedimento), chegando a uma amostra final de 361 participantes (55% mulheres, média de idade: 34,1, idade padrão: 11,1).<sup>3</sup>

### 2.1.2. Procedimento

Os participantes leram as instruções explicando a tarefa e receberam definições das seis emoções, retiradas de dicionários padrão, que eles deveriam avaliar no estudo:

- *Raiva* — Um forte sentimento de desprazer, indignação, aborrecimento e hostilidade.
- *Embaraço* — Um sentimento de dolorosa autoconsciência, desconforto, constrangimento e humilhação.
- *Medo* — Um sentimento desagradável de ansiedade ou apreensão causado pela presença ou expectativa de perigo.
- *Culpa* — Consciência de ter agido errado ou falhado em uma obrigação, acompanhado de sentimentos de remorso e arrependimento.
- *Tristeza* — Um sentimento de infelicidade, sofrimento e amargura.

---

<sup>3</sup> Notamos que nos dois experimentos relatamos todas as medidas, manipulações e exclusões nos experimentos. Além disso, todos os critérios de exclusão e análises estatísticas foram planejados antes de se realizar a análise e nenhuma coleta de dados prosseguiu depois que as análises começaram.

- *Vergonha* — Uma emoção negativa, dolorosa, que combina sentimentos de desonra, indignidade, angústia e humilhação, e é causada pela consciência de ter feito algo errado o idiota.

Para garantir que os participantes entenderam os termos de emoção, eles receberam seis definições e lhes foi pedido que selecionassem a emoção correspondente. Os participantes que responderam incorretamente a essas questões foram excluídos das análises.

Em seguida os participantes leram cinco versões de um cenário, cada uma com um nível diferente de conhecimento sobre o evento focal, apresentadas em ordem aleatória. Depois de cada versão, os participantes usaram barras deslizantes para relatar com que força experimentariam cada uma das seis emoções em uma escala de 0 a 100, e com que probabilidade e intensidade exibiriam 16 reações físicas, também em uma escala de 0 a 100. Essas reações foram retiradas de pesquisas anteriores sobre emoções e incluíam: evitar o olhar, empalidecer, enrubescer, cerrar os dentes, cobrir ou tocar o rosto, chorar, encarar, baixar a cabeça, rir nervosamente, arregalar os olhos, erguer as sobrancelhas, sorrir encabulado, encolher os ombros, gaguejar e tremer (ver [Ekman, 2003](#); [Keltner & Buswell, 1997](#); [Tracy & Robins, 2004](#)).

Em seguida os participantes relataram informações demográficas básicas e responderam uma verificação simples de atenção, disfarçada como uma pergunta final que lhes pedia para reportar quais emoções de uma lista de 20 eles estavam experimentando. Havia uma instrução embutida no lembrete que lhes pedia para selecionar “somente a opção entusiástica” para mostrar que eles estavam prestando atenção; os participantes que não a selecionavam eram excluídos das análises. Em seguida recebiam uma explicação e o pagamento.

Os seguintes três cenários foram projetados para provocar emoções autoconscientes:

- *Zombaria* — Os participantes imaginavam que estavam conversando com um amigo e zombando do problema de fala de uma amiga mútua quando a pessoa que era motivo da zombaria aparecia inesperadamente.
- *Flatulência* — Os participantes se imaginavam flatulando sonoramente em uma sala de conferências lotada.
- *Falsificação* — Os participantes imaginavam que estavam apresentando uma solicitação desonesta de reembolso para um clube, que é descoberta por seu companheiro.

Os dois cenários a seguir foram projetados para provocar raiva ou tristeza:

- *Rompimento* — Os participantes imaginavam que seu parceiro romântico ia romper com eles, já que ouviram um correio de voz deixado por um amigo do parceiro.
- *Sabotagem* — Os participantes imaginavam que ouviram por acaso um conhecido tentando sabotar seu relacionamento com um bom amigo.

Cada participante leu cinco diferentes versões de um determinado cenário, cada uma com um nível diferente de conhecimento sobre o evento focal. Como os participantes assumiam o papel do ator nos cenários autoconscientes, mas de observador nos cenários de emoções básicas, os níveis de conhecimento nos dois tipos de cenários não eram idênticos. Os cenários de emoções autoconscientes foram apresentados com cada um dos níveis de conhecimento abaixo:

1. *Conhecimento privado 1* — Somente o participante sabe sobre sua transgressão.

2. *Conhecimento privado 2* — Um observador está consciente da transgressão, mas o participante não percebe isso.
3. *Conhecimento compartilhado 1* — O participante sabe que um observador sabe sobre a transgressão, mas o observador não está consciente que o participante sabe disso.
4. *Conhecimento compartilhado 2* — O participante sabe que um observador sabe da transgressão, o observador sabe que o participante sabe disso, mas o observador não está consciente que o participante sabe que o observador sabe disso.
5. *Conhecimento comum* — A transgressão é conhecida em comum (ou seja, o participante sabe que o observador sabe disso, sabe que o observador sabe que ele sabe, e assim indefinidamente).

Os cenários de emoções básicas foram apresentados com cada um dos seguintes níveis de conhecimento:

1. *Sem conhecimento* — O participante não está consciente do evento focal.
2. *Conhecimento privado 1* — O participante está consciente do evento focal, mas a outra pessoa não percebe isso.
3. *Conhecimento privado 2* — O participante está consciente do evento focal e a outra pessoa sabe disso, mas o participante não sabe que a outra pessoa está consciente de que ele sabe.
4. *Conhecimento compartilhado* — O participante está consciente do evento focal, e o participante sabe que a outra pessoa está consciente que ele sabe disso, mas a outra pessoa não está consciente que ele sabe disso.
5. *Conhecimento comum* — O evento focal é conhecido em comum (isto é, a outra pessoa sabe que o participante sabe disso, sabe que o participante sabe que ela sabe, e assim indefinidamente).

Os níveis de conhecimento não foram descritos nessas formas esquemáticas (que seriam difíceis para os participantes controlarem), mas em descrições concretas do que os atores olharam, notaram ou deixaram de notar. Por exemplo, na condição de Conhecimento Compartilhado 2 do cenário de Zombaria, em que o participante imagina que está zombando do problema de fala de uma amiga, Lisa, numa conversa com um amigo, Robert, a descrição diz: “Robert ri da piada interna exatamente quando Lisa aparece saindo de uma loja por perto. Ele escuta você zombando do problema de fala dela e a expressão facial dela muda. Você tenta não fazer contato visual com ela olhando rapidamente na direção contrária; no entanto, você percebe que ela conseguiu apanhar você no ato de desviar o olhar. Entretanto, nesse ponto, você já havia virado o suficiente para não parecer para ela que você notou que ela viu você.” O texto completo de todas as frases consta dos Materiais Suplementares 2.

### 2.1.3. Análise de dados

Embora vergonha, embaraço e culpa sejam emoções distintas, elas se sobrepõem o suficiente para que as pessoas frequentemente as misturem ao relatar seus sentimentos (Keltner & Buswell, 1996; Tangney et al., 1996; Tangney & Tracy, 2012). Em uma análise preliminar, confirmamos que as avaliações dos participantes para embaraço, vergonha e culpa estavam altamente correlacionadas em todas as condições de conhecimento de todos os três cenários autoconscientes. Assim, combinamos essas emoções autoconscientes e tiramos a médias das avaliações (as confiabilidades dessas 15 notas de escala de emoções autoconscientes variaram de  $\alpha = .791$  até  $\alpha = .948$ ). Da mesma maneira, as reações físicas auto-reportadas associadas às emoções autoconscientes — ruborizar, cobrir ou tocar o rosto, abaixar a cabeça, riso nervoso,

sorriso encabulado, ombros abaixados e gagueira<sup>4</sup> — também estavam altamente correlacionadas; assim, mais uma vez calculamos as médias dessas avaliações (as confiabilidades dessas 15 avaliações de escala variaram de  $\alpha = .682$  até  $\alpha = .850$ ).<sup>5</sup> Finalmente, para distinguir emoções autoconscientes de negatividade geral, subtraímos um compósito das avaliações das outras emoções negativas da medida combinada de emoção autoconsciente.

## 2.2. Resultados

### 2.2.1. Cenários de emoções autoconscientes

A [Figura 1](#) apresenta as avaliações para emoções autoconscientes ([Fig. 1a](#)) e reações físicas associadas ([Fig. 1b](#)) em todos os níveis de conhecimento em cada um dos três cenários autoconscientes. ANOVAs de medidas repetidas (ver [Tabela 1](#)) confirmaram que as emoções autoconscientes relatadas diferiam significativamente nos níveis de conhecimento nos cenários de falsificação, zombaria e flatulência. Testes planejados pareados em todos os níveis de conhecimento adjacentes confirmaram que em todos os três cenários, as avaliações para emoções autoconscientes eram mais altas na condição de conhecimento comum do que nas condições de conhecimento compartilhado (todas  $ps < .001$ ), e mais altas em cada uma das condições de conhecimento compartilhado do que nas condições de conhecimento privado (todas  $ps < .001$ ). ANOVAs de medidas repetidas também confirmaram que as reações físicas relatadas associadas a emoções autoconscientes diferiam significativamente nos níveis de conhecimento em falsificação,  $F(4, 256) = 23.12, p < .001, \eta^2_p = .27$ , zombaria,  $F(4, 308) = 36.95, p < .001, \eta^2_p = .32$ , e cenários de flatulência,  $F(4, 276) = 51.08, p < .001, \eta^2_p = 0.43$ . Testes pareados planejados em níveis adjacentes de conhecimento confirmaram em todos os três cenários que as avaliações dessas reações eram mais altas nas condições de conhecimento comum do que em qualquer condição de conhecimento compartilhado (todas  $ps < .005$ , exceto  $p = .016$  para conhecimento compartilhado 1 X conhecimento privado 1 no cenário de Falsificação).

---

<sup>4</sup> Não se incluiu evitar o olhar nestes compostos, embora com frequência esteja associado às emoções autoconscientes porque evitar ou fazer contato visual pode alterar o nível de conhecimento predominante.

<sup>5</sup> As avaliações médias para cada condição e cenário de conhecimento constam dos Materiais Suplementares 1.

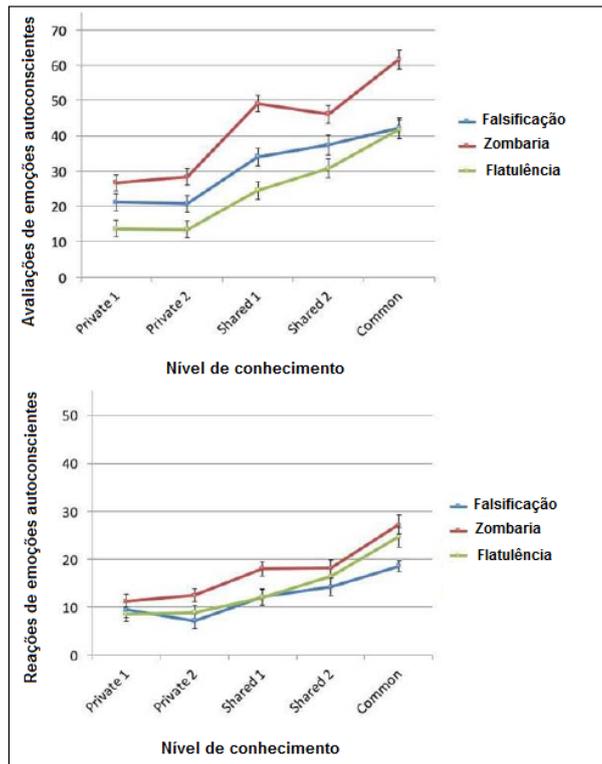
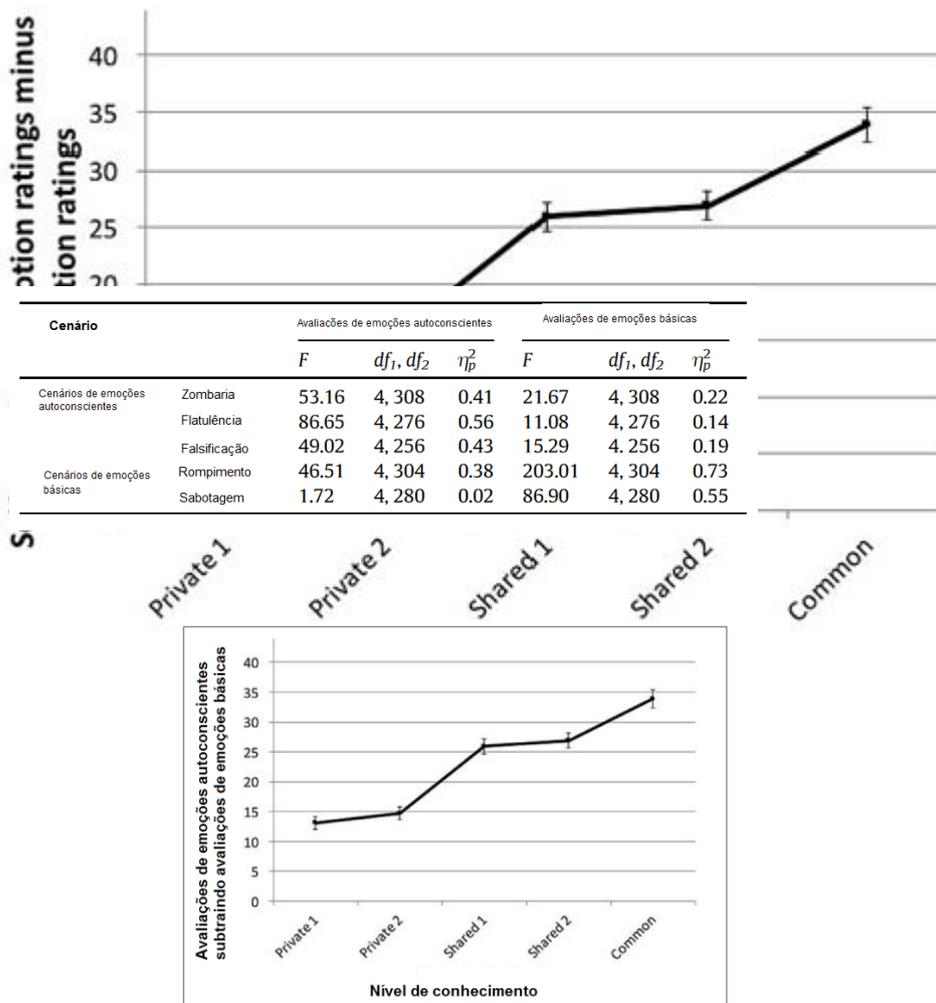


Fig. 1. Avaliações de emoções autoconscientes (1a, no topo) e probabilidade de exibir seus correspondentes físicos (1b, na parte de baixo) nos três cenários de emoções autoconscientes, em diferentes níveis de conhecimento. Cada linha na Fig. 1a representa as avaliações médias combinadas de embaraço, vergonha e culpa nos respectivos cenários. Cada linha na Fig. 1b representa as avaliações médias combinadas de sete reações físicas associadas a essas emoções em seus respectivos cenários. As barras de erro representam erro padrão.



e avaliações de

Fig. 2. Avaliações médias para emoções autoconscientes subtraindo-se as avaliações médias para emoções básicas, colapsadas em todos os três cenários de emoções autoconscientes no Experimento 1. Barras de erro representam erro padrão.

Em seguida, examinamos se a sensibilidade aos níveis de conhecimento é característica das emoções autoconscientes negativas em particular, como prediz a teoria da coordenação, ou se ela se aplica a quaisquer emoções negativas. Analisamos se os níveis de conhecimento afetaram as avaliações dos participantes para emoções de raiva, medo e tristeza e calculamos a média para formar um índice único de afeto negativo. Essa medida de afeto negativo geral mostrou um ligeiro aumento nos níveis de conhecimento em todos os três cenários, mas criticamente, os efeitos dos níveis de conhecimento foram menores (indo de  $\eta^2_p = .14$  até  $\eta^2_p = .22$ ) do que para as emoções autoconscientes (indo de  $\eta^2_p = .41$  até  $\eta^2_p = .43$ ).<sup>6</sup>

Em seguida testamos se os níveis cenário conhecimento moldam as emoções autoconscientes, especificamente, depois de subtrair o afeto negativo geral. Subtraímos o composto de efeito negativo (descrito acima) do composto de emoções autoconscientes para cada participante dentro de cada nível de conhecimento, e depois colapsamos essas notas diferentes em todos os três cenários. A Fig. 2 mostra as emoções autoconscientes distintas dos participantes por nível de conhecimento, e mais uma vez, o padrão apóia a hipótese de coordenação.

Especificamente, encontramos uma diferença geral significativa,  $F(4, 848) = 107.21, p < .001, \eta^2_p = .336$ , com conhecimento comum maior do que qualquer uma das condições de conhecimento compartilhado (ambas  $ps < .001$ ), e cada condição de conhecimento compartilhado maior do que qualquer uma das condições de conhecimento privado (todas  $ps < .001$ ).

### 2.2.2. Cenários de emoções básicas

<sup>6</sup> Ver estatísticas inferenciais na Tabela 1, e estatística descritiva em Materiais Suplementares 1.

Em seguida, analisamos se as emoções não autoconscientes de raiva e tristeza são sensíveis a níveis de conhecimento nos cenários destinados a produzir essas emoções. A Fig. 3 mostra avaliações de tristeza no cenário de rompimento e raiva no cenário de sabotagem nos níveis de conhecimento (Fig. 3a), e as respectivas respostas físicas de chorar e cerrar os dentes (Fig. 3b). ANOVAs de medidas repetidas (ver Tabela 1) revelaram diferenças significativas nos níveis de conhecimento para avaliações de tristeza no cenário de rompimento, e raiva no cenário de sabotagem, bem como choro no cenário de rompimento,  $F(4, 304) = 72.33, p < .001, \eta^2p = .49$ , e cerrar os dentes no cenário de sabotagem,  $F(4, 280) = 34.42, p < .001, \eta^2p = .33$ .

Testes posteriores mostraram que essas diferenças entre níveis de conhecimento basicamente se deviam a uma grande diferença entre as respostas na condição sem conhecimento e as outras quatro condições (todas  $ps < .001$ ). As avaliações para tristeza no cenário de rompimento não diferiam significativamente entre as quatro condições em que o participante sabia do evento (todas  $ps > .19$ ); no entanto, a probabilidade relatada de choro era significativamente mais baixa na condição de conhecimento comum do que em qualquer uma das condições de conhecimento compartilhado ou privado (todas  $ps < .001$ ). As avaliações para raiva no cenário de sabotagem não diferiram entre as condições de conhecimento privado e compartilhado (todas  $ps > .64$ ); no entanto, as avaliações de raiva foram significativamente maiores na condição de conhecimento comum do que na condição de conhecimento compartilhado ( $p = .004$ ). Inesperadamente, os participantes relataram uma probabilidade mais baixa de cerrar os dentes na condição conhecimento privado 2 do que na condição de conhecimento comum ( $p = .015$ ) e na condição de conhecimento privado 1 ( $p = .009$ ).

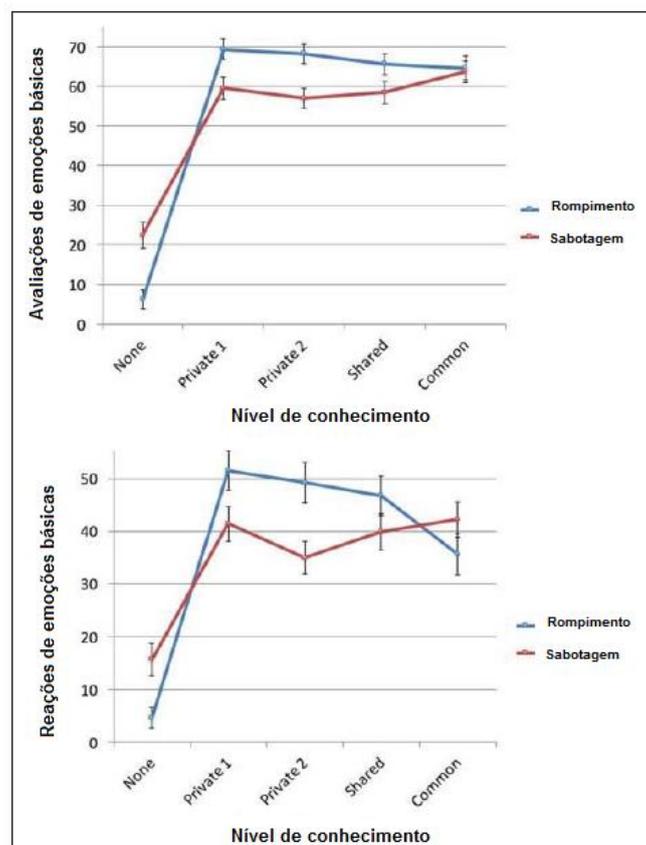


Fig. 3. Avaliações para tristeza no cenário de rompimento e raiva no cenário de sabotagem (3a, no topo), e avaliações da probabilidade de choro no cenário de rompimento e cerrar os dentes no cenário de sabotagem (3b, na parte de baixo), nos diferentes níveis de conhecimento. As barras de erro representam erro padrão.

### 2.3. Discussão

Os resultados do Experimento 1 replicam a bem conhecida descoberta que as emoções autoconscientes são sentidas com maior intensidade e têm mais probabilidade de serem expressas em reações físicas características quando uma transgressão é testemunhada por espectadores (conhecimento compartilhado), em comparação com quando só é conhecida privadamente (conhecimento privado). Mais importante, os resultados mostram que as pessoas relatam que sentiriam emoções autoconscientes mais intensas quando sua transgressão é de conhecimento comum do que quando é conhecimento compartilhado. Essa observação confirma uma previsão distintiva da hipótese de coordenação para emoções autoconscientes. Ainda mais, esses efeitos eram particulares às emoções autoconscientes. Os cenários que provocam raiva e tristeza mostraram um padrão diferente: a raiva e a tristeza esperadas foram evocadas principalmente por se descobrir privadamente sobre o evento desencadeador, com pouco ou nenhum aumento à medida que os níveis de conhecimento aumentavam para além do conhecimento privado. Dessa forma, as pessoas esperam que essas emoções não autoconscientes sejam evocadas simplesmente por saber de um evento desencadeante, ao passo que as emoções autoconscientes seguem diferentes níveis de conhecimento compartilhados com uma audiência sobre o evento.

### 3. Experimento 2

O Experimento 2 traz o fenômeno para dentro do laboratório a fim de testar se o embaraço se intensifica com o conhecimento de uma situação comprometedora real em oposição a uma situação comprometedora meramente imaginada. Tanto a sabedoria convencional como a pesquisa de laboratório (Hofmann, Moscovitch, & Kim, 2006) sugerem que, para a maior parte das pessoas, cantar diante de uma audiência é um desafio social com potencial de embaraço significativo, presumivelmente porque o cantor está demandando a atenção da audiência com a promessa implícita de entretê-la mas com uma alta probabilidade de fracasso (particularmente porque a autoconsciência sobre o próprio desempenho compromete o desempenho).

Neste experimento os participantes cantaram para um painel de quatro jurados em uma sala separada através de uma transmissão bidirecional em vídeo e depois relataram com que força sentiram diferentes emoções. Durante sua apresentação, os participantes olhavam um vídeo em estilo de karaokê com a letra das músicas inseridas em um visor maior de vídeo mostrando o que os participantes pensavam que era uma transmissão ao vivo do júri na outra sala, forçando-os a olhar o júri durante toda a sua apresentação (ver Fig. 4a e b). Na realidade, os participantes estavam vendo um vídeo pré-gravado de um júri em vez de um júri ao vivo. Imediatamente depois de sua apresentação, os participantes avaliavam como se sentiram para uma bateria de emoções. As principais medidas dependentes eram as avaliações em bruto dos participantes para embaraço, bem como uma medida mais refinada de embaraço diferenciado, em que subtraímos o afeto negativo geral (avaliações médias para outras emoções negativas) das avaliações de embaraço dos participantes.

Em um desenho entre sujeitos, manipulamos dois fatores relativos a níveis de conhecimento. A primeira manipulação era que diziam ao participante que: (1) os jurados sabiam que havia transmissão bidirecional de vídeo ao vivo e sabiam que o participante tinha consciência que eles sabiam disso (condição de conhecimento comum); ou, (2) os jurados falsamente acreditavam que estavam assistindo uma transmissão unidirecional de vídeo e pensavam que o participante não tinha consciência da presença deles (condição de conhecimento

compartilhado). A segunda manipulação era que os jurados no vídeo estavam olhando direto na câmara (condição de contato visual; Fig. 4a) ou estavam olhando para o lado (condição sem contato visual; Fig. 4b). Essas manipulações destinavam-se a variar as crenças dos participantes sobre o conhecimento comum e contato visual como um sinal psicofísico para conhecimento comum.



Fig. 4. Capturas de tela representativas dos vídeos que os participantes viam nas condições de contato visual (4a, topo) e sem contato visual (4b, parte de baixo) no Experimento 2. A tela preta com a letra da música no canto inferior esquerdo era um vídeo integrado que mostrava a letra da música do karaokê.

A teoria da coordenação prevê maior embaraço quando um participante pensa que sua cantoria ruim é de conhecimento comum. Na condição de conhecimento compartilhado, um participante está cantando mal e sabe que há desconhecidos assistindo, mas pelo menos os desconhecidos pensam que ele está cantando só para si, e além disso, esses desconhecidos ainda poderiam fingir que não viram o que aconteceu para ajudar o participante a não fazer um papel tão feio. No entanto, na condição de conhecimento comum, o participante não pode se consolar com essas considerações: o fato que os jurados viram sua cantoria ruim em um vídeo bidirecional e que isso era de conhecimento comum significa que nenhuma das partes pode fingir desconhecimento da outra. A má cantoria do participante é um fato inevitável e evidente perante ele e os jurados.

Sendo assim, comparando o conhecimento compartilhado e o conhecimento comum, podemos testar se a emoção autoconsciente de embaraço é sensível a níveis diferenciados de conhecimento. Se as pessoas se sentem embaraçadas principalmente em resposta ao que as outras pessoas sabem sobre seus fiascos, então elas deveriam se sentir igualmente embaraçadas quando os jurados ouvem sua cantoria ruim, seja que os jurados saibam ou não que o participante sabe que eles estão observando. Porém, se as pessoas se sentem extremamente embaraçadas quando seus fiascos são de conhecimento comum, então os

participantes ficarão mais embaraçados quando todos tiverem consciência que os jurados ouviram sua cantoria ruim.

Dessa forma, este experimento testa a previsão da teoria de coordenação sobre como o embaraço será afetado por diferentes níveis de conhecimento (conhecimento comum X compartilhado) e por um potencial sinal psicofísico de conhecimento comum (contato visual X nenhum contato visual) em um design fatorial  $2 \times 2$  entre sujeitos.

### 3.1. Método

#### 3.1.1. Participantes

Com base em uma análise de poder para detectar tamanhos médios de afeto, recrutamos 126 participantes (na faixa etária de 18 a 75 anos) do grupo de estudo da Universidade de Harvard, que receberam dez dólares créditos de curso pela participação. Como a desconfiança do participante em experimentos envolvendo engano é uma preocupação específica, isso foi avaliado usando-se um debriefing em funil estruturado e excluíram-se os participantes que demonstraram altos níveis de desconfiança ( $n = 10$ ), ou que realmente não cantaram a canção de karaokê ( $n = 3$ ). Obtivemos uma amostra final de  $N = 113$  (57% mulheres, média de idade = 28,1, idade padrão = 13,3).<sup>7</sup>

#### 3.1.2. Procedimento

Os participantes se inscreveram em um estudo descrito como um experimento sobre como o anonimato afeta as avaliações dos julgadores sobre o desempenho de uma pessoa. O estudo afirmava que os participantes seriam indicados ou para cantar em karaokê ou para julgar a apresentação de outra pessoa. Ao chegarem, os participantes ouviam que estávamos estudando como as relações entre os jurados e cantores afeta as avaliações dos jurados — “uma importante questão não respondida na idade dos reality shows na TV” — e que em cada estudo recrutamos um painel de jurados e um cantor. Em seguida todos os participantes foram informados que tinham sido indicados para o papel de cantores e que quatro outros participantes indicados para o papel de jurados já estavam em uma sala do outro lado do saguão. Foi dito aos participantes que eles cantariam em uma apresentação de karaokê que o júri assistiria em uma transmissão bidirecional de vídeo ao vivo e que essa disposição de duas salas era necessária para garantir que o ambiente fosse o mesmo quando os juízes acreditassem que eram anônimos e quando não acreditassem.

Os participantes indicados para a condição de conhecimento comum receberam as seguintes instruções: “Você foi selecionado para a condição na qual os jurados sabem que você está consciente da presença deles. Você conseguirá vê-los pela transmissão bidirecional de vídeo durante sua apresentação e eles sabem que você consegue vê-los pela transmissão bidirecional de vídeo. Em outras palavras, tanto você como os juízes sabem que vocês estão se assistindo em uma transmissão bidirecional de vídeo durante sua apresentação.” Assim, nesta condição de conhecimento comum, os participantes acreditavam que todos sabiam que todos sabiam (e assim por diante) que os juízes estavam assistindo o participante (e vice-versa) através de uma transmissão bidirecional de vídeo ao vivo.

Os participantes indicados para a condição de conhecimento compartilhado receberam as seguintes instruções: “Você foi selecionado para a condição anônima, em que os jurados pensam que você não está consciente da presença deles. No entanto, para manter as

---

<sup>7</sup> As perguntas utilizadas no questionário de funil constam dos Materiais Suplementares 4, e as estatísticas descritivas e tamanhos de amostras constam dos Materiais Suplementares 5. Todos os critérios de exclusão foram especificados antes da coleta de dados, e todas as decisões de exclusão foram feitas antes de à análise estatística dos dados.

apresentações iguais em todas as condições, nós realmente sempre contamos aos cantores sobre os juízes. Além disso, você conseguirá vê-los na transmissão bidirecional de vídeo durante sua apresentação, mas eles não sabem disso. Em outras palavras, embora você consiga ver os juízes através da transmissão, eles pensam que você não sabe que eles estão lá, e que eles estão simplesmente assistindo uma transmissão unidirecional de vídeo de sua apresentação.” Assim, nessa condição de conhecimento compartilhado, os participantes acreditavam que os jurados pensavam que estavam assistindo uma transmissão unidirecional de vídeo e que o participante não tinha consciência da presença deles (embora o participante, na realidade, sabia que eles estavam lá e conseguia vê-los pela transmissão bidirecional ao vivo).

Foi dito aos participantes nas duas condições que, depois de terminada a canção, a transmissão em vídeo seria interrompida e eles receberiam uma breve pesquisa.<sup>8</sup>

Em seguida os participantes foram conduzidos a uma pequena sala privada com um grande monitor de computador com câmera de vídeo acoplada, alto-falantes e um microfone para que eles cantassem. O pesquisador dispôs o computador diante do participante e lhe disse que tinha que sair para preparar a transmissão em vídeo na sala do júri. Foi dito aos participantes que assim que o pesquisador preparasse isso, eles veriam uma transmissão de vídeo dos juízes bem como um vídeo de karaokê com a letra das músicas, e foram instruídos a começar sua apresentação assim que a música começasse. Depois de 30 segundos, um vídeo aparecia na tela do computador com a cabeça do pesquisador bem diante da câmera (como se estivesse acabando de ligar a câmera de vídeo na outra sala). Imediatamente depois, a canção *Rolling in the Deep*, de Adele, uma canção popular bem conhecida com um refrão, começava a tocar, e os participantes viam uma grande imagem em vídeo do júri na outra sala, com uma pequena transmissão de karaokê incorporada no canto inferior esquerdo que mostrava a letra da canção (ver Fig. 4). Na condição de *contato visual*, o vídeo dos juízes os mostrava olhando diretamente para a câmera (Fig. 4a); na condição *sem contato visual*, o vídeo os mostrava olhando para outro lado, como se estivessem olhando para uma tela a uns 60 cm ao lado da câmera (Fig. 4b).

Depois de terminada a música, o vídeo parava e aparecia uma pesquisa na tela do computador, pedindo aos participantes para avaliarem a intensidade com que experimentaram diferentes emoções em uma escala de 1 a 5 em uma Escala de Afeto Positivo e Afeto Negativo (PANAS) levemente modificada. A PANAS inclui 10 itens, compreendendo uma escala de afeto negativo e 10 itens compreendendo uma escala de afeto positivo (ver Watson et al., 1988). Essa escala foi usada para disfarçar nosso interesse no embaraço, inserindo-a em uma bateria maior de avaliações de emoções. No entanto, a PANAS não pergunta especificamente sobre embaraço; assim, um dos itens de afeto positivo foi substituído por embaraço.

Em seguida os participantes responderam algumas questões sobre sua experiência com o karaokê e prestaram informações demográficas básicas. Depois saíram da sala, onde se encontraram com o pesquisador e responderam um questionário de funil, que começava com questões vagas sobre o que os participantes achavam que o estudo versava, seguido por questões cada vez mais específicas: se eles achavam que tudo tinha sido explicado claramente e se eles ficaram desconfiados durante o estudo, concluindo com uma pergunta direta se eles acreditavam que os juízes eram reais.<sup>9</sup> Depois disso, os participantes tiveram uma reunião final e receberam seu pagamento.

---

<sup>8</sup> As instruções completas estão apresentadas nos Materiais Suplementares 3.

<sup>9</sup> O protocolo completo das perguntas do questionário em funil está incluído nos Materiais Suplementares 4.

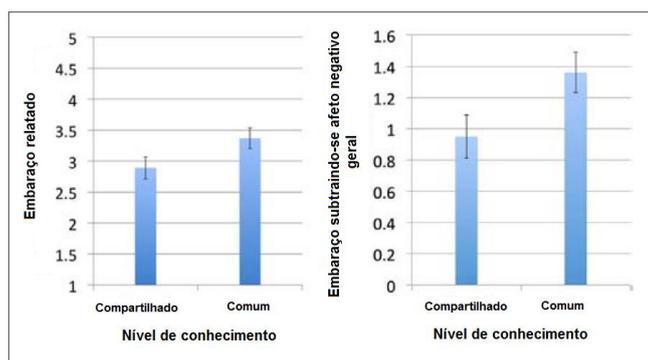


Fig. 5. Média de embaraço relatado (5a, à esquerda), e média de embaraço relatado com o afeto negativo geral subtraído (5b, à direita), nas condições de conhecimento compartilhado e comum no Experimento 2. As barras de erro representam o erro padrão.

### 3.2. Resultados e discussão

A Fig. 5a mostra os resultados. Um ANOVA 2 (conhecimento compartilhado X comum)  $\times$  2 (contato visual X nenhum contato visual) revelou um efeito marginalmente significativo de nível de conhecimento nas avaliações de embaraço,  $F(3, 109) = 3.83, p = .053, \eta^2_p = .03$  (ver Fig. 5a), nenhum efeito de contato visual,  $F(3, 109) = .96, p = .328, \eta^2_p = .01$ , e nenhum efeito de interação,  $F(3, 109) = .15, p = 0.703, \eta^2_p = .00$ . Especificamente, os participantes mostraram uma tendência a maior embaraço na condição de conhecimento comum do que em conhecimento compartilhado, consistente com a hipótese de coordenação, embora notemos que essa tendência tivesse significância marginal.

Em seguida, assim como no Experimento 1, examinamos o embaraço distinto dos participantes subtraindo outras emoções negativas — neste caso, uma média de 10 itens da escala de afeto negativo de PANAS ( $\alpha = .891$ ). Um ANOVA 2  $\times$  2 revelou um efeito significativo de nível de conhecimento no embaraço menos afeto negativo,  $F(3, 109) = 4.77, p = .031, \eta^2_p = .04$  (ver Fig. 5b), nenhum efeito de contato visual,  $F(3, 109) = 1.82, p = .181, \eta^2_p = .02$ , e nenhum efeito de interação,  $F(3, 109) = .16, p = .691, \eta^2_p = .00$ . Especificamente, os participantes mostraram maior embaraço distinto quando sua apresentação de karaokê era de conhecimento comum do que quando era conhecimento compartilhado.

Em geral, esses resultados fornecem algum suporte inicial para a previsão que as pessoas se sentem mais embaraçadas quando sua apresentação é de conhecimento comum do que quando é conhecimento compartilhado. O efeito foi marginalmente significativo para avaliações brutas de embaraço e atingiu significância para embaraço distintivo. Em contraste, não encontramos evidência que o contato visual, um potencial gerador de conhecimento comum, por si só aumente o embaraço.

## 4. Discussão geral

Em geral, esses experimentos fornecem duas linhas de evidências convergentes que o reconhecimento de uma transgressão com uma audiência evoca sentimentos mais intensos de culpa, embaraço e vergonha do que simplesmente saber que eles sabem disso. Utilizando insights da teoria de jogos, a teoria da coordenação propõe que as pessoas se sentem mais embaraçadas, culpadas ou envergonhadas quando uma transgressão é reconhecida ou, de alguma outra forma, se torna pública por dois motivos: (1) a falha de um ator em redimir-se de uma transgressão transmite mais informação prejudicial quando é conhecimento

compartilhado em comparação com quando uma audiência simplesmente sabe disso; e, (2) as relações humanas são um tipo de jogo de coordenação, que torna uma transgressão mais destrutiva quando é de conhecimento comum em coordenação com conhecimento compartilhado. Essa teoria da coordenação faz a previsão nova que as emoções autoconscientes serão provocadas com mais força pelo conhecimento comum de uma transgressão do que por conhecimento compartilhado.

No Experimento 1, um participante típico relatou que se sentiria mais embaraçado, culpado e envergonhado em cenários hipotéticos quando ele e um observador tivessem conhecimento comum da ofensa do que quando ele e o observador tivessem somente conhecimento compartilhado. Observamos esses efeitos para as avaliações brutas dos participantes de emoções autoconscientes e também para uma medida refinada de emoções autoconscientes distintas, subtraindo afeto negativo geral. Além disso, os efeitos do conhecimento compartilhado X comum eram mais pronunciados para emoções autoconscientes do que para outras emoções negativas, tristeza e raiva.

Uma complicação empírica no Experimento 1 é que os participantes também relataram raiva um tanto maior quando um ato agressivo era de conhecimento comum entre o agressor e a vítima, em vez de meramente conhecimento compartilhado. Essa observação poderia indicar que até algumas das assim chamadas emoções básicas também poderiam ajudar a coordenar as relações, além de outras funções centrais. Por exemplo, a raiva pode envolver elementos de coordenação conforme duas partes renegociam suas expectativas conjuntas da consideração que uma deve à outra (ver [Reed, DeScioli, & Pinker, 2014](#); [Sell, Tooby, & Cosmides, 2009](#)).

No Experimento 2, um participante típico relatou que se sentiu embaraçado com uma apresentação de karaokê diante de um júri quando ele acreditava que os jurados sabiam que ele tinha consciência da presença deles do que quando ele acreditava que os jurados pensavam que eles eram anônimos. Nesse caso, no entanto, as avaliações brutas dos participantes para embaraço atingiram apenas significância marginal. Porém, quando examinamos uma medida refinada de embaraço distinto, depois de subtrair o afeto negativo, descobrimos valores significativamente maiores para conhecimento comum do que para conhecimento compartilhado. Embora esses resultados sejam um tanto misturados, sugerimos que, em geral, eles convergem com Experimento 1 para indicar emoções autoconscientes mais intensas sob conhecimento comum do que conhecimento compartilhado. Em contraste, não encontramos efeitos de contato visual, que teorizamos que seria um potente sinal de conhecimento comum. Abaixo discutimos mais esta observação.

Assim, essas constatações sugerem um adendo importante às teorias anteriores de reputação para emoções autoconscientes: Perceber que outras pessoas sabem sobre a ofensa de alguém pode ser desconfortável, mas perceber que eles sabem que uma pessoa sabe que eles sabem (e assim por diante) pode ser angustiante. Sugerimos que isso acontece porque as emoções autoconscientes ajudam a manter a reputação além de manter as relações coordenando as crenças das pessoas sobre o estado de suas relações.

#### *4.1. Sinais sociais e expiação*

Um indivíduo que viola expectativas sociais sem reconhecer a transgressão está meramente sob uma nuvem de suspeita, e permanece em aberto a possibilidade que ele estabelecerá sua inocência, compensará a vítima ou corrigirá suas atitudes; ou seja, que ele ao menos queira seguir as regras e seja suficientemente competente para fazê-lo. Em contraste, alguém que reconhece a infração mas continua sem pedir desculpas indica que ele não se considera vinculado às regras comunitárias (ou é incapaz de entender ou segui-las) e enxerga pouco valor em pertencer àquela comunidade. Além disso, à medida que a informação sobre a transgressão vaza e se espalha, a credibilidade de qualquer eventual pedido de desculpas

diminui (Sznycer et al., 2015). O agudo desconforto desencadeado por conhecimento compartilhado ou comum de uma transgressão, então, serve para motivar o transgressor a expiar seus atos, tornar tais expiações mais críveis e, assim, proteger sua posição ameaçada na comunidade.

#### *4.2. Relações como jogos de coordenação*

A análise apresentada neste artigo sustenta que as relações sociais podem ser caracterizadas como jogos de coordenação. Dois indivíduos podem se beneficiar se ambos concordarem tacitamente em serem amigos, amantes, vizinhos prestativos, parceiros de negócios, chefe e empregado, ou desconhecidos educados. Esses contratos sociais permitem que ambos os parceiros compartilhem expectativas sobre recursos, papéis e comportamento aceitável sem negociação constante, mal-entendidos e conflitos (Fiske, 1992; Haslam, 2004; Lee & Pinker, 2010; Pinker, 2007). Porém, se duas pessoas assumem modelos diferentes de relação (amante X supervisionado, amigo X chefe, cliente X membro da família) podem surgir conflitos quando uma parte reclama recursos ou pré-requisitos que o outro não está disposto a conceder, resultando em sentimentos de indelicadeza, imposição, insubordinação, autoritarismo, assédio sexual, exploração, ou familiaridade ou formalidade impróprias. Para gerenciar efetivamente diferentes tipos de relações, as emoções autoconscientes ajudam a solucionar esses problemas de coordenação estabelecendo e reparando o conhecimento comum dos tipos de relações.

Os trabalhos anteriores sobre níveis de conhecimento e coordenação se concentraram basicamente no modo como o conhecimento comum pode ajudar os parceiros sociais a obter benefícios mútuos coordenando-se em novas oportunidades (por ex., Chwe, 2001; Schelling, 1960; Thomas et al., 2014). No entanto, o conhecimento comum também pode gerar perdas mútuas desencadeando uma cascata de fuga em que os parceiros estabelecidos tentam adivinhar as futuras intenções uns dos outros, fazendo com que a coordenação se revele.

Dessa forma, o conhecimento comum pode tanto facilitar a coordenação como desmantelá-la. Pode ajudar as pessoas a construir novos relacionamentos mas também pode destruir os existentes, isso porque a coordenação depende de um alicerce instável de expectativas mútuas. Ao tentar se coordenar com um parceiro, uma pessoa precisa fazer o que ela espera que seu parceiro vá fazer, e isso, por sua vez, depende do que ela espera que o parceira espera que ela faça e do que ele espera que ela espera que ele espera que ela faça, e assim por diante. O conhecimento comum pode facilitar a coordenação concentrando as expectativas dos parceiros na mesma opção, de forma que essas expectativas interdependentes se alinhem. O conhecimento comum também pode destruir a coordenação concentrando a atenção dos parceiros em um evento que sugere que seus interesses podem já não estar alinhados, o que pode tornar novamente imprevisíveis as expectativas sobre o comportamento futuro. Essa mudança nas expectativas pode apresentar aos parceiros cooperativos o mesmo dilema que eles enfrentaram quando estabeleceram inicialmente suas relações, e exige que eles tomem medidas para realinhar as expectativas se quiserem voltar a estabelecer o status quo.

O conhecimento compartilhado, em contraste, é menos potente em ambos os casos porque, sem conhecimento comum, os dois parceiros não conseguem ter certeza que suas expectativas do comportamento futuro estejam alinhadas. Sem alinhamento claro das expectativas, o conhecimento compartilhado pode tornar difícil coordenar e estabelecer uma nova relação. No entanto, essa mesma falta de alinhamento também significa que o conhecimento compartilhado de uma transgressão pode deixar os parceiros cooperativos sem espaço de manobra para evitar o processo potencialmente doloroso e caro de renegociar os termos da relação deles. Em resumo, o conhecimento compartilhado permite que os parceiros

cooperativos estabelecidos finjam de modo plausível que nada jamais aconteceu e retomem com mais facilidade o status quo, se assim o quiserem.

#### *4.3. Coordenação por terceiros e emoções autoconscientes*

Além do problema de coordenação dentro de uma relação, também há problemas de coordenação para terceiros fora de uma relação. Se um terceiro que está observando acusa alguém de cometer transgressões ou incompetência, esse observador se coloca em perigo. O acusado pode retaliar com contra-acusações e ataques prejudiciais. No entanto, esses custos se reduzem se outros observadores fazem o mesmo porque a retaliação do infrator agora está dividida entre um grupo de acusadores. Assim, os terceiros que quiserem castigar, excluir ou punir um infrator enfrentam um problema de coordenação: se um indivíduo confronta o infrator sozinho, os custos de retaliação podem ser altos, mas se múltiplos observadores coordenarem suas acusações, a ameaça de retaliação será minimizada.

Assim como em outros problemas de coordenação, os terceiros estão mais equipados para sincronizar sua condenação se têm conhecimento comum sobre a violação do infrator (DeScioli & Kurzban, 2013). Assim, a punição fica menos custosa e, portanto, mais provável quando uma transgressão é de conhecimento comum entre observadores do que se é apenas conhecimento compartilhado.

Da perspectiva do transgressor, isso significa que o conhecimento comum de sua ofensa entre vários observadores oferece uma maior ameaça de punição e assim, maior necessidade de expiar sua infração. As emoções autoconscientes podem responder melhor por ameaça de punição tornando ainda maior o motivo para se desculpar quando uma infração é de conhecimento comum entre os observadores. Ou seja, as emoções autoconscientes podem destinar-se a ajudar as pessoas a gerenciar não apenas suas relações diáticas depois de uma transgressão, mas também as ameaças sociais mais amplas que vêm de múltiplos terceiros que podem ter tomado consciência das indiscrições de um infrator. Se for esse o caso, as emoções autoconscientes também poderiam ter evoluído para rastrear níveis de conhecimento entre terceiros, além de entre uma pessoa e um parceiro social.

Essa extensão da teoria da coordenação das emoções autoconscientes oferece um número de previsões testáveis para pesquisa futura. Mais óbvio, ela prevê que as emoções autoconscientes serão mais intensas quando uma infração for de conhecimento comum entre os observadores — por exemplo, se a transgressão ocorrer em público e for visível para muitas pessoas (que também conseguem se ver observando a transgressão) ou se a infração for registrada em foto ou vídeo que podem ser amplamente vistos em fóruns públicos como no Twitter ou YouTube. Uma segunda previsão é que as emoções autoconscientes das pessoas serão sensíveis a sinais indicando que um número crescente de observadores estão tomando conhecimento comum da ofensa, como quando ouvem outras pessoas focarem sobre a transgressão. Uma terceira previsão é que as emoções autoconscientes motivarão as pessoas a tentar evitar que sua ofensa se torne conhecimento comum, tal como coagir outros a não falarem da ofensa, mesmo se todos os membros da audiência já sabem dela privadamente.

#### *4.4. Conhecimento comum e a exibição de emoções autoconscientes*

Se as emoções autoconscientes evoluíram conforme a lógica dos jogos de coordenação, isso poderia ajudar a explicar outra característica distinta: Elas desencadeiam não apenas exibições faciais localizadas, assim como fazem as emoções básicas, mas também respostas autonômicas evidentes e posturas corporais dinâmicas que alteram a visibilidade do rosto de uma pessoa (Keltner, 1995; Keltner & Buswell, 1997; Lewis, 2000; Tracy & Matsumoto, 2008; Tracy & Robins, 2004). A vergonha e o embaraço podem causar rubor, encolher de ombros, cabeça baixa, evitar olhar de outras pessoas e cobrir o rosto. O orgulho, uma emoção

autoconsciente positiva, está associado a estufar o peito e erguer braços e cabeça. Sugerimos que essas exibições possam ter evoluído em parte para gerar ou evitar conhecimento comum.

Muito do desconforto de ruborizar as faces vem do fato que o rubor é exibido exteriormente e sentido como calor e formigamento interiormente, de forma que a pessoa que cora sabe que os expectadores sabem que ela sabe que eles sabem que ela está vermelha, e assim por diante (ver [Thomas et al., 2014](#)). As posturas corporais abatidas que acompanham a vergonha podem funcionar em dois níveis. Evitando contato visual, o transgressor aparentemente está evitando o conhecimento comum da transgressão. Entretanto, diferentemente de um olhar furtivo de um transgressor que ainda está esperando fugir do conhecimento comum, a postura abatida evidente de uma pessoa envergonhada sinaliza que ela está consciente da necessidade de evitar o conhecimento comum embora já não esteja realmente tentando evitá-lo. Não somente um observador pode perceber que o estado emocional de um transgressor abatido mudou, mas a própria visão que o transgressor tem dos observadores foi obscurecida ou desviada para longe e, fundamentalmente, o observador consegue ver isso e perceber que o transgressor deve estar consciente que o observador consegue ver isso, e assim por diante. (Como apontam [Goffman, 1959, 1978](#), e [Brown & Levinson, 1987](#), não é coincidência que dizemos que uma pessoa que cometeu uma infração social comumente conhecida “ficou com a cara no chão” e que a pessoa que tenta consertar o prejuízo está tentando “salvar a cara”). Um indivíduo abatido pode, assim, tornar conhecimento comum que ele já reconhece a necessidade de mudar seu comportamento, o que pode permitir a todas as partes retomarem o status quo sem ter que se envolver em conflito desnecessário. Portanto, a expressão física das emoções autoconscientes é consistente com a teoria que essas emoções, em uma medida maior do que outras emoções, envolvem raciocínio de estado mental recorrente no serviço de estratégias para negociar os tipos de jogos de coordenação que permeiam a vida social humana.

Uma novidade empírica nesta explicação é que no Experimento 2, contato visual, que deveria, teoricamente, ser um poderoso gerador de conhecimento comum (cada pessoa vê a outra olhando para trás), na realidade não levou a maior embaraço. Essa observação é surpreendente, considerando-se o papel do contato visual na vida social em geral e, especificamente, o fenômeno de evitar o olhar em resposta ao embaraço. A descoberta pode significar que o contato visual direto não afeta o embaraço, nem leva a inferências de conhecimento comum. No entanto, notamos que a atual manipulação do contato visual era relativamente sutil e não convencional; ocorreu através de transmissão por vídeo em que os olhos dos juízes estavam voltados diretamente ao participante ou a um ligeiro ângulo em relação ao participante. Além disso, os juízes na realidade foram filmados antecipadamente, para que seus olhos não pudessem responder dinamicamente às ações do participante, que pode ter parecido como encarando no vazio em vez de prestar atenção ao participante. Pesquisas futuras podem examinar os efeitos do contato visual sob circunstâncias mais naturalistas para testar ainda mais como ele pode estar relacionado ao embaraço e inferências sobre o conhecimento comum.

## **5. Conclusão**

Encerramos sugerindo que o conhecimento comum tem implicações de longo alcance para a psicologia. A diferença de categoria entre o conhecimento compartilhado e comum há tempos é reconhecida por teóricos de jogos ([Rubinstein, 1989](#); [Schelling, 1960](#)), economistas ([Chwe, 2001](#); [Geanakoplos, 1992](#)), filósofos ([Lewis, 1969](#)), linguistas ([Clark, 1996](#); [Clark & Marshall, 1981](#); [Smith, 1982](#)), sociólogos ([Adut, 2008](#)), antropólogos e cientistas políticos ([Cronk & Leech, 2012](#)), teóricos legais ([McAdams & Nadler, 2005](#)) e cientistas computacionais ([Alberucci & Jäger, 2005](#); [Halpern & Moses, 1990](#)). Pesquisas recentes demonstraram que essa distinção também molda a psicologia da cooperação ([Thomas et al.,](#)

2014), ajuda (Thomas et al., 2016), e comunicação (Lee & Pinker, 2010). Os presentes experimentos acrescentam emoções autoconscientes ao conjunto de superfícies psicológicas humanas que são sensíveis às diferenças em categorias entre conhecimento privado, compartilhado e comum.

## Financiamento

Este trabalho teve o apoio da Universidade de Harvard.

## Agradecimentos

Agradecemos a Tiana Press por sua assistência para rodar os experimentos.

## Apêndice A. Dados suplementares

Dados complementares a esta artigo podem ser encontrados online em: <https://doi.org/10.1016/j.evolhumbehav.2017.12.001>.

## Bibliografia

- Adut, A. (2008). *On scandal: Moral disturbances in society, politics, and art*. New York, NY: Cambridge University Press.
- Alberucci, L., & Jäger, G. (2005). About cut elimination for logics of common knowledge. *Ann. Pure Appl. Logic*, 133(1–3), 73–99. <https://doi.org/10.1016/j.apal.2004.10.004>.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323–370. <https://doi.org/10.1037/1089-2680.5.4.323>.
- Beer, J. S., Heerey, E. A., Keltner, D., Scabini, D., & Knight, R. T. (2003). The regulatory function of self-conscious emotions: Insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology*, 85(4), 594–604. <https://doi.org/10.1037/0022-3514.85.4.594>.
- Bosch, J. A., de Gues, E. J. C., Carroll, D., Goedhart, A. D., Anane, L. A., Veldhuizen, J. J., ... Edwards, K. M. (2009). A general enhancement of autonomic and cortisol responses during social evaluative threat. *Psychosomatic Medicine*, 71, 877–885. <https://doi.org/10.1097/PSY.0b013e3181baef05>.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. New York, NY: Cambridge University press.
- Chakroff, A., Thomas, K. A., Haque, O. S., & Young, L. (2015). An indecent proposal: The dual functions of indirect speech. *Cognitive Science*, 39(1), 199–211. <https://doi.org/10.1111/cogs.12145>.
- Chwe, M. S. (2001). *Rational ritual: Culture, coordination, and common knowledge*. Princeton, NJ: Princeton University Press.
- Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, UK: Cambridge University Press.
- Cronk, L., & Leech, B. L. (2012). *Meeting at grand central: Understanding the social and evolutionary roots of cooperation*. Princeton, NJ: Princeton University Press.
- Dalkiran, N. A., Hoffman, M., Paturi, R., Ricketts, D., & Vattani, A. (2012). Conhecimento comum and state-dependent equilibria. *Proceedings of symposium on algorithmic game theory* (pp. 84–95). [https://doi.org/10.1007/978-3-642-33996-7\\_8](https://doi.org/10.1007/978-3-642-33996-7_8).
- DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, 139(2), 477–496. <https://doi.org/10.1037/a0029065>.
- Dijk, C., de Jong, P. J., & Peters, M. L. (2009). The remedial value of blushing in the context of transgressions and mishaps. *Emotion*, 9(2), 287–291. <https://doi.org/10.1037/a0015081>.
- Dijk, C., Koenig, B., Ketelaar, T., & de Jong, P. J. (2011). Saved by the blush: Being trusted despite defecting. *Emotion*, 11(2), 313–319. <https://doi.org/10.1037/a0022774>.
- Ekman, P. (1999). Basic emotions. In T. Dalgleish, & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 45–60). West Sussex, UK: John Wiley & Sons Ltd.
- Ekman, P. (2003). *Emotions revealed*. New York, NY: Times Books.
- Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, 99(4), 689–723. <https://doi.org/10.1037/0033-295X.99.4.689>.
- Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York, NY: W. W. Norton.
- Geanakoplos, J. (1992). Common knowledge. *The Journal of Economic Perspectives*, 6(4), 53–82. <https://doi.org/10.1257/jep.6.4.53>.
- Goffman, E. (1959). *The presentation of self in everyday life*. New York, NY: Anchor Books.
- Goffman, E. (1978). Response cries. *Language*, 54(4), 787–815. <https://doi.org/10.2307/413235>.
- Gruenewald, T. L., Kemeny, M. E., Aziz, N., & Fahey, J. L. (2004). Acute threat to the social self: Shame, social self-esteem, and cortisol activity. *Psychosomatic Medicine*, 66, 915–924. <https://doi.org/10.1097/01.psy.0000143639.61693.ef>.
- Halpern, J. Y., & Moses, Y. (1990). Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3), 549–587. <https://doi.org/10.1145/79147.79161>.
- Haslam, N. (2004). Research on the relational models: An overview. In N. Haslam (Ed.), *Relational models theory: A contemporary overview* (pp. 27–57). Mahwah, NJ: Erlbaum.
- Hofmann, S. G., Moscovitch, D. A., & Kim, H. (2006). Autonomic correlates of social anxiety and embarrassment in shy and non-shy individuals. *International Journal of Psychophysiology*, 61(2), 134–142. <https://doi.org/10.1016/j.ijpsycho.2005.09.003>.
- Holmes, E. A., Arntz, A., & Smucker, M. R. (2007). Imagery rescripting in cognitive behavior therapy: Images, treatment techniques and outcomes. *Journal of Behavior Therapy and Experimental Psychiatry*, 38, 297–305. <https://doi.org/10.1016/j.jbtep.2007.10.007>.
- Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, 68(3), 441–454. <https://doi.org/10.1037/0022-3514.68.3.441>.
- Keltner, D., & Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition and Emotion*, 10(2), 155–171. <https://doi.org/10.1080/026999396380312>.
- Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin*, 122(3), 250–270. <https://doi.org/10.1037/0033-2909.122.3.250>.

- Ketelaar, T., & Au, W. T. (2003). The effects of feelings of guilt on the behavior of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion*, *17*(3), 429–453. <https://doi.org/10.1080/02699930143000662>.
- Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review*, *117*(3), 785–807. <https://doi.org/10.1037/a0019688>.
- Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
- Lewis, M. (2000). Self-conscious emotions: Embarrassment, pride, shame, and guilt. In M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 623–636) (2nd ed.). New York: Guilford.
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*(3), 173–192. <https://doi.org/10.1111/j.1745-6924.2008.0073.x>.
- McAdams, R. H., & Nadler, J. (2005). Testing the focal point theory of legal compliance: The effect of third-party expression in an experimental hawk/dove game. *Journal of Empirical Legal Studies*, *2*(1), 87–123. <https://doi.org/10.1111/j.1740-1461.2005.00032.x>.
- McCullough, M. E. (2008). *Beyond revenge: The evolution of the forgiveness instinct*. San Francisco, CA: Jossey-Bass.
- McCullough, M. E., Kurzban, R., & Tabak, B. A. (2010). Evolved mechanisms for revenge and forgiveness. In P. R. Shaver, & M. Mikulincer (Eds.), *Understanding and reducing aggression, violence, and their consequences* (pp. 221–239). Washington, DC: American Psychological Association.
- Miller, R. S. (1995). On the nature of embarrassability: Shyness, social evaluation, and social skill. *Journal of Personality*, *63*(2), 315–339. <https://doi.org/10.1111/j.1467-6494.1995.tb00812.x>.
- Modigliani, A. (1971). Embarrassment, facework, and eye contact: Testing a theory of embarrassment. *Journal of Personality and Social Psychology*, *17*(1), 15–24. <https://doi.org/10.1037/h0030460>.
- Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York, NY: Viking.
- Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. New York, NY: Viking.
- Pinker, S. (2014). *The sense of style: The thinking person's guide to writing in the 21<sup>st</sup> century*. New York, NY: Penguin.
- Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences*, *105*(3), 833–838. <https://doi.org/10.1073/pnas.0707192105>.
- Reed, L. I., DeScioli, P., & Pinker, S. (2014). The commitment function of angry facial expressions. *Psychological Science*, *25*(8), 1511–1517. <https://doi.org/10.1177/0956797614531027>.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. [https://doi.org/10.1207/S15327957PSPR0504\\_2](https://doi.org/10.1207/S15327957PSPR0504_2).
- Rubin, D. C., & Talarico, J. M. (2009). A comparison of dimensional models of emotion: Evidence from emotions, prototypical events, autobiographical memories, and words. *Memory*, *17*(8), 802–808. <https://doi.org/10.1080/09658210903130764>.
- Rubinstein, A. (1989). The electronic mail game: Strategic behavior under “almost common knowledge.” *The American Economic Review*, *79*(3), 385–391 Retrieved from <http://www.jstor.org/stable/1806851>.
- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819. <https://doi.org/10.1037/0022-3514.76.5.805>.
- Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
- Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, *106*(35), 15073–15078 (doi:10.1073/pnas.0904312106).
- Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge, UK: Cambridge University Press.
- Smith, N. (1982). *Mutual knowledge*. Orlando, FL: Academic Press.
- Smith, R. H., Webster, J. M., Parrot, W. G., & Eyre, H. L. (2002). The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology*, *83*(1), 138–159. <https://doi.org/10.1037/0022-3514.83.1.138>.
- Steckler, C. M., & Tracy, J. L. (2014). The emotional underpinnings of social status. In J. T. Cheng, J. L. Tracy, & C. Anderson (Eds.), *The psychology of social status* (pp. 201–224). New York, NY: Springer.
- Sznycer, D., Schniter, E., Tooby, J., & Cosmides, L. (2015). Regulatory adaptations for delivering information: The case of confession. *Evolution and Human Behavior*, *36*, 44–51. <https://doi.org/10.1016/j.evolhumbehav.2014.08.008>.
- Sznycer, D., Tooby, J., Cosmides, L., Porat, R., Shalvi, S., & Halperin, E. (2016). Shame closely tracks the threat of devaluation by others, even across cultures. *Proceedings of the National Academy of Sciences* (Advance online publication) <https://doi.org/10.1073/pnas.1514699113>.
- Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, *20*(6), 1256–1269. <https://doi.org/10.1037/0022-3514.70.6.1256>.
- Tangney, J. P., & Tracy, J. L. (2012). Self-conscious emotions. In M. Leary, & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 446–478) (2nd ed.). New York, NY: Guilford.
- Thomas, K. A., De Freitas, J., DeScioli, P., & Pinker, S. (2016). Recursive mentalizing and common knowledge in the bystander effect. *Journal of Experimental Psychology: General* (Advance online publication) <https://doi.org/10.1037/xge0000153>.
- Thomas, K. A., DeScioli, P., Haque, O. S., & Pinker, S. (2014). The psychology of coordination and common knowledge. *Journal of Personality and Social Psychology*, *107*(4), 657–676. <https://doi.org/10.1037/a0037037>.
- Tooby, J., & Cosmides, L. (1996). Friendship and the banker's paradox: Other pathways to the evolution of adaptations for altruism. In W. G. Runciman, J. Maynard Smith, & R. I. M. Dunbar (Eds.), *Evolution of social behavior patterns in primates and man. Proceedings of the British Academy*, *88*. (pp. 119–143) Retrieved from <http://www.cep.ucsb.edu/papers/friendship.pdf>.
- Tracy, J. L., & Matsumoto, D. (2008). The spontaneous expression of pride and shame: Evidence for biologically innate nonverbal displays. *Proceedings of the National Academy of Sciences*, *105*(33), 11655–11660. <https://doi.org/10.1073/pnas.0802686105>.
- Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory*, *15*(2), 103–125. [https://doi.org/10.1207/s15327965pli1502\\_01](https://doi.org/10.1207/s15327965pli1502_01).
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*(1), 35–57 Retrieved from <http://www.jstor.org/stable/2822435>.
- Utikal, V. (2012). A fault confessed is half redressed—Confessions and punishment. *Journal of Economic Behavior & Organization*, *81*(1), 314–327. <https://doi.org/10.1016/j.jebo.2011.11.001>.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>.
- Weiner, B., Graham, S., Peter, O., & Zmuidinas, M. (1991). Public confession and forgiveness. *Journal of Personality*, *59*(2), 281–312. <https://doi.org/10.1111/j.1467-6494.1991.tb00777.x>.
- Willer, R., Kuwabara, K., & Macy, M. W. (2009). The false enforcement of unpopular norms. *American Journal of Sociology*, *115*(2), 451–490. <https://doi.org/10.1086/599250>.
- Wolf, S. T., Cohen, T. R., Panter, A. T., & Insko, C. A. (2010). Shame proneness and guilt proneness: Toward the further understanding of reactions to public and private transgressions. *Self and Identity*, *9*, 337–362. <https://doi.org/10.1080/15298860903106843>.