

Deixe sua avaliação: o viés nas avaliações formais em plataformas de serviços colaborativos após uma falha

Louise Helene Gonçalves Foernges

2022

1. INTRODUÇÃO

Nos últimos anos, o mundo tem visto um movimento crescente em direção ao uso sustentável dos recursos (LINDBLOM & LINDBLOM, 2017). Aliado a isto, os avanços tecnológicos permitiram o surgimento da “economia compartilhada” (MALHOTRA & VAN ALSTYNE, 2014), também chamados serviços colaborativos. Neste tipo de serviço, indivíduos têm acesso temporário a um bem ou serviço sem transferência de propriedade (BELK, 2014b). Normalmente, uma plataforma online conecta os usuários que estão dispostos a prestar um serviço ou compartilhar um recurso com usuários que procuram aquele serviço ou recurso (BENOIT et al., 2017). Uber, Lyft e Airbnb, por exemplo, são plataformas que se destacam neste segmento, que continua a crescer em número de usuários e relevância econômica, com uma projeção de valer \$335 bilhões até 2025 (FORBES MAGAZINE, 2019).

Ressalta-se que algumas interações de serviços colaborativos são muito interpessoais (BELK, 2014a), ou seja, interações que envolvem um alto nível de contato pessoal, como quando alguém abre sua casa para um estranho no Airbnb ou se torna hóspede na casa de um estranho (BRIDGES & VÁSQUEZ, 2016). Como essas interações interpessoais ocorrem entre estranhos, as plataformas de serviços colaborativos comumente empregam mecanismos de reputação/feedback (BELK, 2014b; BRIDGES & VÁSQUEZ, 2016; HOFMANN et al., 2017; BENOIT et al., 2017). Esses sistemas de feedback autorregulados (geralmente na forma de classificações, ou seja, notas dadas pelos usuários aos prestadores e *reviews*) ajudam a minimizar riscos, desencorajar o mau comportamento e criar confiança entre os pares em serviços colaborativos (BELK, 2014b; BRIDGES, VÁSQUEZ, 2016; HOFMANN, et al., 2017).

Evidências sugerem que os sistemas de reputação/feedback em serviços colaborativos não são totalmente confiáveis e o viés de feedback ocorre com frequência, pois os pares tendem a evitar avaliações/*reviews* negativos nesse contexto (ZERVAS et al., 2015; FRADKIN et al., 2015; BRIDGES & VÁSQUEZ, 2016). A subnotificação de experiências negativas, por sua vez, tem sido associada à natureza pessoal dos serviços colaborativos, onde presumivelmente as “normas sociais” estão sendo seguidas (BRIDGES & VÁSQUEZ, 2016; ZERVAS et al., 2015). Devido à importância dos mecanismos de reputação/feedback para

ajudar os usuários a decidir quem é confiável e mitigar que indivíduos ajam puramente em interesse próprio (BELK, 2014b; HAMARI et al., 2016; BRIDGES & VÁSQUEZ, 2016), os mecanismos de feedback se tornam especialmente importantes na ocorrência de falhas de serviço.

Propomos que a antecipação da culpa, aspecto ligado à harmonia social e à empatia (MICELI, 1992), está por trás do viés positivo de feedback em avaliações formais em serviços colaborativos, ao invés de reciprocidade ou medo de retaliação – variáveis frequentemente encontradas na literatura como mediadoras para o viés de feedback em sistemas de avaliação mútuos (onde tanto o provedor, quanto o usuário são capazes de avaliar um ao outro) (DELLAROCAS & WOOD, 2008; RESNICK & ZECKHAUSER, 2002). Além do efeito do tipo de avaliação (formal ou informal) no feedback por meio da antecipação de culpa, também abordamos o impacto que diferentes tipos de falhas de serviço (competência e calor humano) e a nota geral do prestador (alta, baixa) têm nesse efeito.

Para investigar nossas hipóteses sobre aspectos que potencialmente interferem e explicam o viés positivo de feedback em serviços colaborativos, realizamos dois estudos experimentais. Usamos uma abordagem semelhante a Zervas et al. (2015), com uma condição controle onde normas sociais não afetariam a avaliação. Comparamos o feedback na forma de avaliação (ou seja, classificação/nota) e gorjeta, em um tipo de avaliação formal (ou seja, a tradicional avaliação do aplicativo - ou “no sistema”) e informal (ou seja, para um amigo - ou “fora do sistema”), que serviu como condição de controle, em um contexto de transporte sob demanda análogo ao Uber.

Segundo Guyader (2018), faltam pesquisas sobre como os pares (usuários e provedores) integram aspectos da troca de mercado e normas pró-sociais em suas práticas e interações entre si. O autor acrescenta que investigar melhor as práticas de consumo colaborativo beneficiariam a pesquisa de serviços. Nosso estudo mostra como certos aspectos pró-sociais podem ser determinantes para a validade do feedback em serviços colaborativos, contribuindo assim para o desenvolvimento da teoria sobre o assunto. Além disso, descobrimos que os mecanismos explicativos usuais (medo de retaliação, reciprocidade) para

o viés de feedback não se aplicam ao contexto de serviços colaborativos, mas a antecipação de culpa sim, explica esse efeito.

2. O IMPACTO DAS AVALIAÇÕES FORMAIS NO FEEDBACK DOS USUÁRIOS

Em mercados de serviços colaborativos, muitas vezes os sistemas de feedback e reputação são empregados para mitigar as ações do usuário em interesse próprio e gerar confiança entre eles (RESNICK et al., 2000; JØSANG, ISMAIL & BOYD, 2007; BELK, 2014b; HAMARI et al., 2016). Além disso, a existência de tais sistemas visa motivar os indivíduos a se comportarem de forma responsável (HOFMANN et al., 2017). Esses sistemas permitem que os pares avaliem e/ou revisem uns aos outros e, às vezes, exibam uma nota geral (média das classificações recebidas).

Há evidências que sugerem a existência de viés de avaliação nesses mecanismos de feedback (ZERVAS et al., 2015; FRADKIN et al., 2015; BRIDGES & VÁSQUEZ, 2016). De acordo com Zervas et al. (2015) é possível que “efeitos sociológicos” levem as pessoas a serem mais diplomáticas em suas avaliações em serviços de consumo colaborativo. Da mesma forma, Bridges e Vásquez (2016) argumentam que ao revisar experiências menos do que positivas, os usuários preferem deixar comentários neutros em vez de negativos. É quase como se os usuários seguissem uma “norma” implícita estabelecida ao deixar comentários. “Normas que regem a comunicação e a interação são estabelecidas em um determinado espaço online pelos membros da comunidade que interagem uns com os outros naquele espaço” (BRIDGES & VÁSQUEZ, 2016, p.14). Por tanto, propomos a seguinte hipótese:

H1: Na ocorrência de uma determinada falha de serviço, ao avaliar formalmente (no sistema) o prestador (vs. informalmente/fora do sistema), os usuários darão uma avaliação mais positiva.

Outra forma de feedback incorporada a serviços colaborativos, como o Uber, é a gorjeta. Segundo Azar (2009), a gorjeta costuma ser uma demonstração de gratidão pela qualidade do serviço. Da mesma forma, Lynn e McCall (2000) argumentam que a gorjeta é outra forma de um cliente exercer o controle de qualidade sobre o serviço, funcionando, de fato, como um mecanismo semelhante às avaliações. Os autores argumentam que, consistente com a teoria das motivações de equidade, existe uma correlação positiva entre as avaliações

de serviço e o tamanho da gorjeta. No entanto, como em um contexto de serviços colaborativos as avaliações podem potencialmente ter consequências de longo prazo para o provedor, enquanto a quantidade de gorjeta não prejudicaria permanentemente a subsistência do provedor, propomos que as gorjetas sejam um método de feedback menos enviesado do que as avaliações.

H2: Na ocorrência de uma determinada falha de serviço, a gorjeta será uma forma de feedback menos enviesada do que as avaliações (na forma de classificações/notas).

3. ANTECIPAÇÃO DE CULPA COMO MEDIADOR

Como apontam Bridges e Vásquez (2016), fatores socioculturais como polidez e cortesia podem ser uma das razões para o viés positivo dos feedbacks no Airbnb. Da mesma forma, Zervas et al. (2015) argumentam que em serviços colaborativos (por exemplo, Airbnb), fatores sociológicos muitas vezes levam os usuários a serem mais diplomáticos em suas avaliações do que em serviços “profissionais” (por exemplo, hotéis). A proximidade com o provedor parece levar os usuários a serem mais empáticos e evitar deixar avaliações negativas para não serem ‘indelicados’ (BRIDGES & VÁSQUEZ, 2016).

Vale ressaltar que, embora a literatura existente aponte que o viés de feedback pode ser explicado pela reciprocidade ou medo de retaliação (DELLAROCAS & WOOD, 2008; RESNICK et al., 2000 RESNICK & ZECKHAUSER, 2002; BOLTON et al., 2013; FRADKIN et al., 2015), devido à natureza pessoal dos serviços colaborativos, a aparente existência de normas sociais neste contexto (o que parece não ocorrer em serviços tradicionais) e o fato de que a maioria dos sistemas de avaliação nesses serviços são agora duplo-cegos (ou seja, usuários e provedores não têm acesso à avaliação um do outro até que ambas tenham sido enviadas à plataforma, acreditamos que reciprocidade e medo de retaliação não explicam o viés positivo nas avaliações de serviços colaborativos. Em vez disso, sentimentos de antecipação de culpa explicarão o efeito do tipo de avaliação (formal, informal) no feedback (na forma de classificação/nota).

Para Baumeister, Stillwell e Heatherton (1994, p.243) “a culpa é algo que acontece entre as pessoas e não apenas dentro delas. Ou seja, a culpa é um fenômeno interpessoal que está funcional e causalmente ligado às relações comunitárias entre as pessoas”. Os autores

acrescentam ainda que o sentimento de culpa vem de uma antecipação -ou do sentimento real- do sofrimento do outro. Portanto, a antecipação da culpa é responsável por um indivíduo realizar ou evitar determinadas ações. Em consonância com isso, Steenhaut e Kenhove (2006) argumentam que a antecipação da culpa funciona como um mecanismo para interromper determinado comportamento ou para controlar a ação.

A antecipação da culpa pode ser despertada pelo pensamento de uma transgressão ou fracasso, que as pessoas tendem a evitar, e uma motivação para “cumprir solicitações comportamentais que as ajudarão a evitar futuro sentimento de culpa” (LINDSEY, YUN & HILL, 2007, p.468). Assim, propomos que a antecipação de culpa é o mecanismo que explica o efeito do tipo de avaliação na classificação/nota, em vez do medo de retaliação ou reciprocidade.

H3: Na ocorrência de uma determinada falha de serviço, o efeito do tipo de avaliação no feedback (classificação/nota) será mediado pela antecipação de culpa.

4. FALHA DE SERVIÇO

Uma das principais razões pelas quais existem sistemas de feedback em serviços colaborativos é garantir que as falhas sejam detectadas e relatadas (RESNICK et al., 2000). Uma falha de serviço ocorre quando há um problema na entrega de um serviço e ele não atende às expectativas do cliente (HESS JR., GANESAN & KLEIN, 2003). Falhas no atendimento são inevitáveis e frequentemente provocam sentimentos e reações negativas no cliente (SMITH & BOLTON, 1998), muitas vezes levando a um sentimento de confiança violada (WANG & HUFF, 2007; BASSO & PIZZUTTI, 2016).

Segundo Mattila (2001), serviços que envolvem um alto grau de contato humano são particularmente propensos a falhas. Três tipos de falhas são comumente encontrados na literatura: moralidade, competência e calor humano. Uma falha de competência ocorre quando o vendedor/prestador de serviços não possui as habilidades e/ou recursos para realizar uma tarefa, deixando de satisfazer o cliente (WANG & HUFF, 2007). De acordo com Kirmani et al. (2017) o calor humano inclui traços de ser sociável, brincalhão, feliz e engraçado. Já uma falha de calor humano ocorre quando esses traços faltam no provedor (ser hostil, frio, anti-social etc). Segundo os autores, as informações sobre moralidade são mais difíceis para os

consumidores observarem e avaliarem ao usar provedores de serviços. Portanto, decidimos testar apenas as falhas relacionadas à competência e ao calor humano.

Seguindo a lógica de Kirmani et al. (2017), que os usuários tendem a valorizar mais a competência do que os traços de calor humano em um provedor de serviços, propomos que: controlando a severidade da falha, uma falha de competência será percebida como aquela que mais compromete a qualidade do serviço principal, levando a menor tendência de perdão do que outros tipos de falha (tal como, calor humano). Assim, uma falha de competência não acarretará diferença significativa na classificação/nota dada ao prestador e quantidade de gorjeta entre os tipos de avaliação (será menos enviesada); enquanto uma falha de calor humano acarretará maiores médias de classificação/nota e quantidade de gorjeta no tipo formal de avaliação (vs. controle - informal).

H4: Na ocorrência de uma determinada falha de serviço, o tipo de falha moderará o efeito do tipo de avaliação na classificação/nota dada pelos usuários, de modo que **a)** para uma falha de serviço de competência, as classificações nos tipos formal e informal de avaliação não serão significativamente diferentes, enquanto **b)** para falha de calor humano, as classificações serão mais altas no sistema formal (vs sistema informal).

H5: Na ocorrência de uma determinada falha de serviço, o tipo de falha moderará o efeito do tipo de avaliação na gorjeta, de modo que **a)** para uma falha de serviço de competência, as gorjetas nos tipos formal e informal de avaliação não serão significativamente diferentes, enquanto **b)** para uma falha de calor humano, a quantidade de gorjeta será maior no sistema formal (vs sistema informal).

5. ESTUDO 1

Design e participantes

O primeiro experimento foi um design fatorial 2 (tipo de falha: competência e calor humano) x 2 (tipo de avaliação: formal -no sistema, controle -informal) entre-sujeitos com atribuição aleatória. O objetivo do Estudo 1 foi testar se as avaliações formais são, de fato, positivamente enviesadas, se a reciprocidade ou o medo de retaliação são mediadores para o efeito do tipo de avaliação (formal, informal) na classificação/nota dada pelo usuário e o papel

moderador do tipo de falha. Para este estudo, uma amostra de 239 participantes foi recrutada por meio do serviço Mechanical Turk (MTurk) da Amazon. A maioria dos participantes (65,3%) declarou ter entre 18-34 anos e 53,1% eram do sexo feminino. Do total da amostra (N = 239), 10 (4,2%) participantes se declararam deficientes.

Procedimento

Aos participantes foi apresentado um vídeo de aproximadamente 1 minuto silenciado e legendado. Cada vinheta continha cenas idênticas mostrando uma corrida de transporte sob demanda (como Uber). No entanto, os vídeos foram legendados de formas diferentes, representando diferentes interações (diálogos) entre passageiro e motorista, introduzindo a manipulação de tipo de falha (competência ou calor humano). Na sequência do vídeo, foi apresentado aos participantes um texto apresentando a manipulação para o tipo de avaliação. Na condição formal (no sistema), os participantes foram informados de que estariam avaliando o motorista no aplicativo do serviço, em uma escala de 1 a 5. Na condição controle (fora do sistema), os participantes foram informados que estariam avaliando sua experiência a um amigo, também em uma escala de 1 a 5.

Medidas

A avaliação foi medida com uma escala deslizante que os usuários foram solicitados a usar para indicar como classificariam o motorista no vídeo, em uma escala que varia de 1 a 5. A gorjeta foi medida com uma escala deslizante que os participantes foram solicitados a usar para indicar a quantidade de gorjeta que estariam dispostos a dar ao motorista, em uma escala que variava de 0% a 25% do preço total da corrida. As seguintes variáveis foram todas medidas em escalas Likert de 7 pontos.

Para verificar as manipulações por tipo de falha (competência e calor humano) foram utilizadas as escalas fornecidas por Leach, Ellemers e Barreto (2007), cada uma composta por 4 itens. O medo de retaliação foi medido com uma pergunta, adaptada de Kudish, Fortunato e Smith (2006): “Tenho medo de sofrer consequências negativas se der um feedback honesto a esse motorista”. A reciprocidade também foi medida com um item: “Avaliei o motorista de acordo com a forma como esperava que o motorista me avaliasse como passageiro”. A gravidade percebida da falha foi medida usando uma escala de 3 itens (por exemplo, “Se o

inconveniente durante o passeio no vídeo estivesse realmente acontecendo com você, você consideraria isso”), de Weun, Beatty e Jones (2004). Todas as medidas foram em escalas de 7 pontos.

Classificação/Nota dada pelo usuário

Para testar os efeitos principais e interativos entre o tipo de falha e o tipo de avaliação na classificação, realizamos um teste ANCOVA controlando a severidade percebida da falha. A análise ANCOVA mostrou um efeito direto significativo do tipo de avaliação (fomal, informal) na classificação dada pelos usuários. Nenhum efeito direto significativo do tipo de falha na classificação foi encontrado ($p > 0,05$). Ao considerar cada falha individualmente, os resultados indicaram que não há diferença nas médias de classificação entre as condições, portanto, H4a foi suportado, mas H4b não. O teste inicial da ANCOVA revelou que os diferentes tipos de avaliação afetaram a maneira como o usuário avaliou sua experiência com o motorista. Em apoio a H1, o tipo de avaliação formal (no sistema) rendeu uma classificação significativamente ($p < 0,001$) mais alta ($M = 2,88$) do que a condição de controle ($M = 2,46$). No entanto, o teste mostrou que para classificação/nota, o efeito interativo do tipo de falha e tipo de avaliação não foi significativo ($F(1,234) = 0,515$, $p > .05$, $\eta^2 p = 0,002$).

Reciprocidade e medo de retaliação

Os resultados da análise de mediação confirmaram que não existe mediação de reciprocidade ou medo de retaliação entre o tipo de avaliação e a classificação/nota dada pelo usuário.

Gorjeta

Realizamos uma ANCOVA para testar os efeitos principais e interativos entre o tipo de falha de serviço (competência, calor humano) e o tipo de avaliação (formal, informal) na gorjeta. O teste não indicou efeito principal do tipo de falha de serviço na gorjeta ($p > ,05$), nem do tipo de avaliação na gorjeta $F(1,234) = 0,200$, $p > ,05$, $\eta^2 p = 0,001$). Essa constatação, juntamente com a constatação de que há uma diferença significativa entre os tipos de avaliação na classificação/nota dada pelo usu[ário, dá suporte à H2, sugerindo que as gorjetas tendem a ser menos tendenciosas do que as classificações. Os resultados também mostraram

um efeito interativo significativo entre tipo de falha de serviço e tipo de avaliação na gorjeta ($F(1,234) = 4,963, p < 0,05, \eta^2 p = 0,021$).

Para investigar o efeito interativo do tipo de avaliação e do tipo de falha na gorjeta, foi realizada uma análise Spotlight. Ao considerar os tipos de falha de competência e calor humano individualmente, os resultados indicam não haver diferença significativa ($p > 0,05$) nas médias de gorjeta entre as condições, portanto H5a é suportado, mas H5b não.

Discussão

Os resultados do Estudo 1 indicam fortes efeitos principais do tipo de avaliação (formal, informal) na classificação/nota dada pelo usuário, mas nenhum efeito interativo. Os resultados revelaram que a média de avaliação foi significativamente maior no tipo formal de avaliação quando comparado à condição controle (informal), dando suporte a H1.

Em H4a, propusemos que para uma falha de serviço de competência, a média de classificação permaneceria inalterada entre os tipos de condições de avaliação, o que foi confirmado pelos nossos resultados. Em H4b propusemos que, no caso de uma falha de calor humano, as médias de classificação seriam maiores (ou seja, mais tendenciosas) em um tipo de avaliação formal (vs informal), o que não ocorreu, refutando, portanto, H4b. O mesmo ocorreu para H5a e H5b, e como as médias de gorjeta foram inalteradas entre os tipos de condições de avaliação, H5a é confirmado, mas não H5b.

Também descobrimos que em um tipo formal de avaliação as médias de gorjeta foram significativamente maiores para uma falha de competência do que para uma falha de calor humano. Uma falha de calor humano geralmente depende das habilidades sociais de um indivíduo (Kirmani et al., 2017), enquanto uma falha de competência geralmente é específica da situação. Weiner et al. (1976) argumentam que a estabilidade de uma causa determina mudanças de expectativa. De acordo com os autores, se as condições causais forem percebidas como passíveis de mudança, não se espera que o resultado atual ocorra novamente. Esse resultado pode indicar que os pares percebem uma falha de calor humano como mais estável (não uma falha pontual) e mais sob controle do provedor, portanto, usando a gorjeta (uma forma de feedback sem potencial de afetar a subsistência do provedor) para exercer

controle de qualidade sobre o serviço prestado. Em seguida, exploramos uma variável que também pode ser percebida como um indicativo de estabilidade da falha.

6. NOTA GERAL DO PRESTADOR

As notas gerais dos prestadores (normalmente medidas em ‘estrelas’), fazem parte dos mecanismos de feedback/reputação em serviços de transporte sob demanda (Uber) e compartilhamento de quartos (Airbnb). Esses mecanismos permitem que os pares estabeleçam uma reputação com base nas avaliações de desempenho de outros pares (WEBER, 2014). De acordo com Bridges e Vásquez (2016), essas pontuações são uma ferramenta importante, pois servem como uma sugestão para o comportamento passado dos colegas e são baseadas na experiência pessoal de outros usuários. Os autores apontam que vários estudos mostram que os indivíduos levam em consideração as avaliações on-line antes de tomar decisões. Usar as pontuações dos colegas como uma pista do comportamento passado está de acordo com a teoria da atribuição. Segundo Weiner (1972), a atribuição de estabilidade é a expectativa de que a causa de um evento permanecerá estável e não flutuante ao longo do tempo. O autor argumenta que (p. 556-557) “se se espera que as condições (a presença ou ausência de causas) permaneçam as mesmas, espera-se que o(s) resultado(s) experimentado(s) no passado se repitam”. Em outras palavras, a nota do driver ajuda os usuários a atribuir estabilidade à falha experimentada.

Como aponta a literatura, é possível que as avaliações/revisões anteriores sirvam de indício de comportamento passado para usuários/clientes de serviços colaborativos (BRIDGES & VÁSQUEZ, 2016). Propomos que quando a nota do motorista for alta, os usuários tenderão a acreditar que a transgressão não é um problema recorrente, pois a nota serve como sugestão de comportamento passado adequado, portanto, ao avaliar formalmente o motorista, uma nota alta do motorista levará a resultados mais positivos. (ou seja, tendenciosa) do que na avaliação informal. No entanto, e mais importante, quando o motorista tem uma nota baixa, apontando para um comportamento passado inadequado, espera-se que os meios de classificação permaneçam os mesmos entre os tipos de condições de avaliação, ou seja, os consumidores avaliarão o motorista de forma mais objetiva no ambiente formal sistema. Ou seja, nessa situação, os passageiros podem achar que as falhas causadas pelo

provedor são recorrentes e se sentirem menos obrigados a dar um feedback positivo (preconceituoso) sobre o motorista (afinal, é provável que ele tenha se comportado mal no passado como bem e não parece merecer uma classificação “falsa” boa desta vez).

H6: A nota geral do motorista moderará o efeito do tipo de avaliação na classificação dos usuários, de modo que na condição de nota alta do motorista, a média de classificação será significativamente maior no tipo formal de avaliação do que no controle (informal), enquanto quando a nota do motorista é baixa, as médias de classificação permanecerão inalteradas entre os tipos de condições de avaliação.

7. ESTUDO 2

Design e participantes

O estudo 2 foi um design fatorial 2 (competência, calor humano) x 2 (nota geral do prestador: alto, baixo) x 2 (tipo de avaliação: formal -no sistema, controle -informal) entre sujeitos com atribuição aleatória. Semelhante ao Estudo 1, o Estudo 2 também foi realizado online, via Mechanical Turk. A amostra final incluiu 349 indivíduos com idade média de 33,9 anos onde 53% eram do sexo masculino. Do total da amostra, 10 (2,9%) participantes se declararam deficientes.

Procedimento

A manipulação de vídeo para o tipo de falha de serviço foi a mesma do Estudo 1. Após o vídeo, uma imagem apresentou a manipulação para nota geral do prestador. A imagem mostrava uma captura de tela de um smartphone mostrando o aplicativo depois que o passageiro solicitou o carro pelo app. Na condição de alta nota do motorista, uma nota de 4,98 foi mostrada sob o nome do piloto com uma estrela ao lado. Na condição de baixa nota do motorista, uma nota de 3,29 foi mostrada. Após foi introduzida a mesma manipulação do tipo de avaliação (formal, informal) do Estudo 1.

Medidas

No Estudo 2, introduzimos as escalas para nota geral do motorista e antecipação de culpa. A nota geral do motorista foi medida com um item: “Você considera a classificação

atual (nota) do motorista:”, com três opções de resposta: “alta”, “baixa” e “média”. Os itens para antecipação de culpa foram adaptados de Basil et al. (2006). Os dois itens foram avaliados em escalas de 7 pontos.

Principais resultados do estudo

Classificação/Nota dada pelo usuário

Para testar os efeitos principais e interativos entre o tipo de falha de serviço e o tipo de avaliação na classificação, realizamos um teste ANCOVA que não mostrou efeitos principais significativos de tipo de falha de serviço ou nota geral do motorista ($p > .05$), mas na condição formal de tipo de avaliação, as notas foram, como no Estudo 1, significativamente ($p < .05$) maiores ($M = 3,25$) do que na condição informal ($M = 3,01$), confirmando, portanto, H1. Além disso, o teste ANCOVA revelou um efeito interativo da nota geral do motorista e do tipo de avaliação (formal, informal) na classificação/nota dada pelo usuário ($F(1, 340) = 7,035$, $p < 0,05$, $\eta^2_p = 0,020$).

Uma análise do Spotlight confirmou que ocorre um efeito interativo entre a nota geral do prestador e o tipo de avaliação. Os resultados do teste mostram que quando a nota geral do prestador é alta, há uma diferença significativa ($p < .05$) nas médias de classificação entre a condição formal ($M = 3,40$) e a condição de controle ($M = 2,84$). No entanto, quando a nota geral do prestador é baixa, não há diferença significativa ($p > .05$) nas médias de classificação entre os tipos de condições de avaliação, confirmando, portanto, H6.

Mediação da Antecipação de Culpa

No Estudo 2, incluímos a antecipação de culpa como possível mediador para o efeito do tipo de avaliação (formal, informal) na classificação/nota dada pelo usuário. Os resultados da análise de mediação revelaram um caminho significativo do tipo de avaliação para a antecipação da culpa ($b = 0,22$ se = $0,10$, $t = 2,13$, $p < 0,05$, com intervalos de confiança (IC) entre $0,01$ e $0,42$) e antecipação de culpa para classificação ($b = 0,25$ se = $0,02$, $t = 9,52$, $p < .05$, com intervalos de confiança (IC) entre $0,20$ e $0,30$), enquanto o efeito direto entre tipo de avaliação e classificação desaparece, tornando-se não significativo ($p > .05$), revelando, portanto, uma mediação plena. Um teste ANOVA revelou que em um tipo formal ($M = 4,25$)

de avaliação, a média de antecipação de culpa foi significativamente ($F(1, 347) = 4,568$, $p < 0,05$) maior do que na condição controle (informal) ($M = 2,81$).

Discussão

Os resultados do Estudo 2 confirmaram a maioria dos resultados do Estudo 1, revelando uma diferença significativa nas médias de classificação/nota dada pelo usuário ao prestador entre os tipos de avaliação formal e controle (informal), dando suporte adicional para H1. A antecipação da culpa revelou-se mediadora desse efeito, oferecendo suporte a H3. Contrariamente à nossa hipótese, verificamos que as falhas de calor humano não apresentam diferença significativa nas médias de classificação ou gorjeta entre os tipos de condições de avaliação, refutando H5a e H5b. Da mesma forma, na ocorrência de falha de competência, não foram encontradas diferenças nas médias de classificação/nota ou gorjeta entre os tipos de condições de avaliação, em apoio a H4a e H4b. Esse resultado sugere que o tipo de falha (ou seja, competência e calor humano) não interfere no viés de feedback. Por fim, os resultados indicam que a nota geral do prestador modera o efeito do tipo de avaliação na classificação/nota dada pelo usuário. Na condição de nota alta do motorista, as classificações foram significativamente diferentes entre os tipos de condições de avaliação. Os resultados revelaram que na condição de alta nota do condutor, as médias de classificação foram maiores no tipo formal de avaliação do que na condição controle (informal), dando suporte a H6. Na condição de baixa nota do motorista, não foi encontrada diferença na média de classificação entre os tipos de condições de avaliação. Portanto, ao avaliar o motorista formalmente, os usuários deram uma classificação mais alta (mais tendenciosa) ao motorista quando a nota do motorista era alta. No tipo de avaliação informal (controle), as classificações médias não foram significativamente diferentes entre a nota baixa e alta do motorista.

8. CONSIDERAÇÕES FINAIS

Ambos os estudos revelaram que, ao avaliar formalmente um provedor, as classificações/notas dadas pelos usuários foram significativamente mais altas (mais positivamente enviesadas) quando comparadas a um tipo de avaliação informal e menos “comprometedor” (ou seja, para um amigo). No Estudo 1, confirmamos que o medo de retaliação e reciprocidade de fato não são mediadores para o efeito do tipo de avaliação na

classificação/nota. Especulamos que isso pode ser devido às características do provedor não ser um profissional e ter uma interação mais próxima com o usuário, o que pode suscitar preocupação empática. É possível que o conhecimento de que a classificação/nota dada possa ter consequências negativas para o provedor possa suscitar empatia dos usuários, ao avaliar formalmente os pares.

No Estudo 2, encontramos a antecipação da culpa como mediadora do efeito do tipo de avaliação (formal, informal) na classificação/nota. Isso está de acordo com a ideia de que os usuários desses serviços tendem, devido à natureza interpessoal destes serviços, a atenuar feedbacks negativos para evitar prejudicar o prestador (BRIDGES & VÁSQUEZ, 2016). Isso está de acordo com Miceli (1992), que argumenta que sentimentos de culpa podem estar ligados à empatia e ao perdão.

Em nenhum dos estudos a quantidade de gorjeta teve uma diferença significativa entre as condições formais e de controle (informais) de avaliação, enquanto a classificação/nota teve em ambos os estudos. Uma possível explicação para isso é, novamente, a antecipação da culpa. É possível que, diferentemente das classificações - que os usuários acreditam que possam ter consequências de longo prazo para o motorista (como demonstrado por várias respostas à nossa pergunta aberta), a culpa não seja provocada ao dar gorjeta, pois isso não afetaria gravemente a subsistência do provedor. Em consonância com isso, Lynn (2009) descobriu que recompensar o serviço era um motivador maior para dar gorjeta do que evitar a culpa. Portanto, a quantidade de gorjetas pode ser uma forma mais confiável de feedback da qualidade do serviço do que as classificações/notas dadas pelos usuários em serviços colaborativos.

No Estudo 2, os resultados indicaram que uma nota alta do motorista leva a um viés positivo no feedback ao avaliar formalmente o provedor. Os resultados revelaram que as médias de classificação/nota foram significativamente maiores no tipo formal de avaliação do que na condição controle (informal), quando os respondentes foram submetidos à condição de motorista com nota geral alta; para os respondentes na condição de motorista com nota geral baixa, a média de classificação/nota dada não foi significativamente diferente entre as condições de avaliação do tipo formal ou informal. Isso é consistente com a teoria da

atribuição que postula que pistas de comportamento passado (a nota do motorista neste caso) são frequentemente usadas como preditores de comportamento futuro (WEINER et al., 1976).

Mecanismos de feedback/reputação pouco confiáveis podem levar a incidentes potencialmente prejudiciais à imagem da empresa (BENOIT et al., 2017). Tendo em vista que a maioria dos serviços colaborativos são autorregulados, entender o que leva os usuários a dar feedbacks enviesados é fundamental para a manutenção da qualidade dos serviços colaborativos e, conseqüentemente, da imagem da empresa. Ao compreender melhor esses aspectos, gestores poderão criar formas de mitigar feedbacks tendenciosos e criar incentivos, como treinar usuários e provedores sobre como fazer bom uso das ferramentas de avaliação.

9. LIMITAÇÕES E SUGESTÕES PARA PESQUISAS FUTURAS

Esta pesquisa não foi isenta de limitações. Primeiramente, ambos os estudos foram realizados online e na mesma plataforma (MTurk). Para obter maior validade externa, sugere-se a realização de mais estudos em outras plataformas (por exemplo, Prolific), ou em campo. Em segundo lugar, estudamos apenas um contexto de serviço colaborativo. Para permitir maior generalização dos resultados, recomenda-se que estudos futuros se concentrem em outros contextos de consumo colaborativo, como a hospitalidade (por exemplo, Airbnb).

Na questão aberta de nossa pesquisa, muitos participantes afirmaram que preferem não interagir com o provedor, enquanto outros afirmaram sentir-se ofendidos pela falta de interação na condição de falta de calor humano. Estudar como a intimidade ou rapport entre usuários impacta no feedback em serviços colaborativos e como a orientação para o relacionamento (comunal vs. troca) afeta a existência de rapport entre os usuários podem ser temas relevantes para pesquisas futuras.

REFERÊNCIAS

- Azar, O. H. (2010). Tipping motivations and behavior in the US and Israel. *Journal of Applied Social Psychology, 40*(2), 421-457.
- Bardhi, F., & Eckhardt, G. M. (2012). Access-based consumption: The case of car sharing. *Journal of Consumer Research, 39*(4), 881-898.
- Basil, D. Z., Ridgway, N. M., & Basil, M. D. (2006). Anticipation of guilt appeals: The mediating effect of responsibility. *Psychology & Marketing, 23*(12), 1035-1054.

- Basso, K., & Pizzutti, C. (2016). Trust recovery following a double deviation. *Journal of Service Research, 19*(2), 209-223.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Anticipation of guilt: an interpersonal approach. *Psychological bulletin, 115*(2), 243.
- Belk, R. (2014a). Sharing versus pseudo-sharing in Web 2.0. *The Anthropologist, 18*(1), 7-23.
- _____. (2014b). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research, 67*(8), 1595-1600.
- Benoit, S., Baker, T. L., Bolton, R. N., Gruber, T., & Kandampully, J. (2017). A triadic framework for collaborative consumption (CC): Motives, activities and resources & capabilities of actors. *Journal of Business Research, 79*, 219-227.
- Bolton, G., Greiner, B., & Ockenfels, A. (2013). Engineering trust: reciprocity in the production of reputation information. *Management science, 59*(2), 265-285.
- Bridges, J., & Vásquez, C. (2018). If nearly all Airbnb reviews are positive, does that make them meaningless?. *Current Issues in Tourism, 21*(18), 2057-2075.
- Dellarocas, C., & Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science, 54*(3), 460-476.
- Eckhardt, G. M., & Bardhi, F. (2015). The sharing economy isn't about sharing at all. *Harvard business review, 28*.
- Forbes Magazine. The Sharing Economy is Still Growing and Business Should Take Note, 2019. Available at <https://www.forbes.com/sites/forbeslacouncil/2019/03/04/the-sharing-economy-is-still-growing-and-businesses-should-take-note/?sh=2cbc42b54c33>. Accessed November 2nd 2019.
- Fradkin, A., Grewal, E., Holtz, D., & Pearson, M. (2015). Bias and reciprocity in online reviews: Evidence from field experiments on Airbnb. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation* (pp. 641-641).
- Guyader, H. (2018). No one rides for free! Three styles of collaborative consumption. *Journal of Services Marketing, 32*(6), 692-714.
- Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology, 67*(9), 2047-2059.
- Hess Jr, R. L., Ganesan, S., & Klein, N. M. (2003). Service failure and recovery: the impact of relationship factors on customer satisfaction. *Journal of the Academy of Marketing Science, 31*(2), 127-145.
- Hofmann, E., Hartl, B., & Penz, E. (2017). Power versus trust—what matters more in collaborative consumption?. *Journal of Services Marketing, 31*(6), 589-603.
- Jøsang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision support systems, 43*(2), 618-644.
- Kirmani, A., Hamilton, R. W., Thompson, D. V., & Lantzy, S. (2017). Doing well versus doing good: The differential effect of underdog positioning on moral and competent service providers. *Journal of Marketing, 81*(1), 103-117.
- Kudisch, J. D., Fortunato, V. J., & Smith, A. F. (2006). Contextual and individual difference factors predicting individuals' desire to provide upward feedback. *Group & Organization Management, 31*(4), 503-529.
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: the importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of personality and social psychology, 93*(2), 234.

- Lindblom, A., & Lindblom, T. (2017). De-ownership orientation and collaborative consumption during turbulent economic times. *International Journal of Consumer Studies*, 41(4), 431-438.
- Lindsey, L. L. M., Yun, K. A., & Hill, J. B. (2007). Anticipated guilt as motivation to help unknown others: An examination of empathy as a moderator. *Communication Research*, 34(4), 468-480.
- Lynn, M. (2009). Individual differences in self-attributed motives for tipping: Antecedents, consequences, and implications. *International Journal of Hospitality Management*, 28(3), 432-438.
- Lynn, M., & McCall, M. (2000). Gratitude and gratuity: a meta-analysis of research on the service-tipping relationship. *The Journal of Socio-Economics*, 29(2), 203-214.
- Malhotra, A., & Van Alstyne, M. (2014). The dark side of the sharing economy... and how to lighten it. *Communications of the ACM*, 57(11), 24-27.
- Mattila, A. S. (2001). The impact of relationship type on customer loyalty in a context of service failures. *Journal of Service Research*, 4(2), 91-101.
- Miceli, M. (1992). How to make someone feel anticipation of guilty: Strategies of anticipation of guilt inducement and their goals. *Journal for the theory of social behaviour*, 22(1), 81-104.
- Möhlmann, M. (2015). Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. *Journal of Consumer Behaviour*, 14(3), 193-207.
- Resnick, P., & Zeckhauser, R. (2002). Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. In *The Economics of the Internet and E-commerce* (pp. 127-157). Emerald Group Publishing Limited.
- _____, P., Kuwabara, K., Zeckhauser, R., & Friedman, E. (2000). Reputation systems. *Communications of the ACM*, 43(12), 45-48.
- Smith, A. K., & Bolton, R. N. (1998). An experimental investigation of customer reactions to service failure and recovery encounters paradox or peril?. *Journal of service research*, 1(1), 65-81.
- Steenhaut, S., & Van Kenhove, P. (2006). *The Mediating Role of Anticipated Guilt in Consumers' Ethical Decision-Making*. *Journal of Business Ethics*, 69(3), 269-288.
- Wang, S., & Huff, L. C. (2007). Explaining buyers' responses to sellers' violation of trust. *European Journal of Marketing*, 41(9/10), 1033-1052.
- Weiner, B. (1972). Attribution theory, achievement motivation, and the educational process. *Review of educational research*, 42(2), 203-215.
- _____, B., Nierenberg, R., & Goldstein, M. (1976). Social learning (locus of control) versus attributional (causal stability) interpretations of expectancy of success 1. *Journal of Personality*, 44(1), 52-68.
- Weun, S., Beatty, S. E., & Jones, M. A. (2004). The impact of service failure severity on service recovery evaluations and post-recovery relationships. *Journal of Services Marketing*, 18(2), 133-146.
- Zervas, G., Proserpio, D., & Byers, J. (2015). A first look at online reputation on Airbnb, where every stay is above average. SSRN Working Paper 2554500.