Preocupação com a privacidade, confiança e disposição dos consumidores a fornecer informações on-line no contexto do Big Data

Resumo

O surgimento de novas ferramentas tecnológicas, que propiciam técnicas mais avançadas de mineração, tem permitido a interpretação de um extenso volume de dados complexos que compõe o Big Data, por meio do qual as empresas passam a ter acesso às informações geradas pelas atividades que os consumidores desempenham on-line, favorecendo a oferta de produtos e serviços personalizados e a obtenção de vantagens competitivas. Contudo, levanta-se uma problemática acerca da privacidade e da segurança dos consumidores. Assim, este estudo objetiva identificar os fatores que afetam a preocupação dos consumidores com a privacidade e suas consequências na confiança e na disposição a fornecer suas informações on-line. Um modelo foi proposto para mensurar a relação entre essas variáveis, considerando a expertise em utilizar a Internet, o risco percebido e o controle informacional como fatores antecedentes, além dos efeitos moderadores da personalidade, dos valores culturais, do padrão de uso da Internet e do conhecimento sobre o Big Data. Com base nos dados de uma amostra de 864 respondentes, adotando-se técnicas da Modelagem de Equações Estruturais, os resultados mostraram que o risco percebido e o controle informacional determinam a preocupação com a privacidade, que afeta a confiança e a disposição a fornecer informações on-line. Embora não tenham sido identificados efeitos moderadores, alguns aspectos sociais e demográficos influenciaram as variáveis observáveis do modelo. A partir dos resultados, este estudo contribui para que os gestores promovam melhorias nas políticas de privacidade e transparência, e na construção de uma relação de confiança entre consumidores e varejistas on-line.

Palavras-chave: privacidade informacional; consumidor on-line; Big Data.

1 INTRODUÇÃO

Com a crescente informatização da sociedade e a emergência mais recente do fenômeno de convergência tecnológica, as pessoas passaram a fornecer e receber informações em tempo real de qualquer lugar do planeta. O acúmulo desses dados, gerados a partir das atividades desempenhadas on-line, compõe parte do que, atualmente, é conhecido como Big Data, que tem aparecido em relatórios técnico-científicos como uma base tecnológica que propicia uma nova forma de coletar, estocar, consolidar e gerir diversos tipos de dados provenientes de diferentes fontes (JACOBS, 2009; DUMBILL, 2012; IBM, 2012; ULARU et al., 2012; DAVENPORT; DYCHÉ, 2013; KAISLER et al., 2013; WARD; BARKER, 2013; MAURO, GRECO; GRIMALDI, 2015; ALOYSIUS; HOEHLE; VENKATESH, 2016; EASTIN et al., 2016).

Nesse cenário, as empresas viram uma oportunidade para ampliar seus mercados, fazendo o uso do rastreamento de dados para analisar como os indivíduos interagem com suas marcas, localizando e anunciando produtos de acordo com os interesses de seus consumidores. Assim, os mesmos tornam-se fornecedores de dados, de forma voluntária e involuntária (através de *cookies* e históricos de navegação, por exemplo), permitindo que as empresas tracem seus perfis de comportamento e consumo (CHEN; CHIANG; STOREY, 2012; CHEN; ZHANG, 2014; ALOYSIUS; HOEHLE; VENKATESH, 2016).

Com base nessas considerações, se para as empresas o Big Data consiste em uma ferramenta que lhes possibilita orientar suas ações a partir do uso analítico dos dados, para os consumidores, resulta em produtos e serviços personalizados (ADOMAVICIUS; TUZHILIN, 2005; COULDRY; TUROW, 2014). No entanto, ao mesmo tempo, emergem riscos com relação à privacidade informacional dos indivíduos, de modo que as lojas on-line devem se atentar para a segurança dos dados de seus consumidores a fim de evitar reações contrárias pelo mau uso ou má proteção informacional (AWAD; KRISHNAN, 2006).

Diversos estudos enfatizam a necessidade de se adotar medidas para reduzir o nível de preocupação com a privacidade e construir uma forte relação de confiança, uma vez que, para entregar um valor personalizado, elas dependem da disposição dos consumidores a fornecer suas informações pessoais (PHELPS; NOWAK; FERRELL, 2000; SCHOENBACHLER; GORDON, 2002; MYERSCOUGH; LOWE; ALPERT, 2006; LI; SARATHY; XU, 2011). Além disso, destacam a importância de se examinar o comportamento dos consumidores no ambiente virtual, suas características pessoais e suas percepções quanto ao risco e à segurança de seus dados (PHELPS, NOWAK, FERRELL, 2000; HARRIS; HOYE; LIEVENS, 2003; MALHOTRA; KIM; AGARWAL, 2004; TAYLOR; DAVIS; JILLAPALLI, 2009; OSATUYI, 2015).

Assim, o objetivo deste estudo consiste em identificar os fatores que afetam a preocupação do consumidor com a privacidade informacional e suas consequências na confiança e na disposição a fornecer informações on-line, diante das possibilidades atuais de rastreamento digital. Para isso, considerou-se a expertise em utilizar a Internet, o risco percebido e o controle informacional como variáveis antecedentes, além do efeito moderador da personalidade, dos valores culturais, do padrão de uso da Internet e do conhecimento dos consumidores sobre o Big Data.

Este estudo busca apresentar contribuições teóricas à literatura de Marketing ao propor um modelo que avalia alguns fatores que determinam a disposição dos consumidores a fornecer informações on-line, considerando os riscos do atual cenário de rastreamento digital feito pelas empresas para promover produtos e serviços personalizados, além de fornecer aos gestores uma nova perspectiva acerca das implicações do Big Data sobre a construção da relação de confiança da empresa com o consumidor.

2 O CONSUMIDOR ON-LINE E A PRIVACIDADE NO CONTEXTO DO BIG DATA

Diversas empresas que atuam no âmbito on-line passaram a enxergar o Big Data como uma ferramenta que permite identificar, através de técnicas de mineração de dados, padrões implícitos de comportamentos dos indivíduos, estabelecendo uma ponte entre o que as pessoas querem fazer e o que de fato fazem quando interagem umas com as outras (MICHAEL; MILLER, 2013; PENTLAND, 2013). Ao mesmo tempo, lidar com a ética, considerando o consentimento dos indivíduos, a segurança e a privacidade, torna-se um dos principais desafios a serem enfrentados (ZIMMER, 2010; BOYD; CRAWFORD, 2012; MICHAEL; MILLER, 2013; PENTLAND, 2013; ZWITTER, 2014).

Destaca-se que a conceptualização do termo "privacidade" ainda é vaga (MARGULIS, 2003; MOORE, 2008; SOLOVE, 2008). A dificuldade em definir o termo surge mediante os entendimentos e experiências de indivíduos em diferentes contextos sócio-históricos (KASPER, 2005). Para este estudo, será adotada a intepretação da privacidade informacional, cuja fricção ontológica se baseia nas forças que se opõem ao fluxo de informação dentro do espaço digital das TICs (Tecnologias de Informação e Comunicação), que leva o indivíduo a considerar o rastreamento digital como uma forma de agressão à sua identidade pessoal (FLORIDI, 2005).

É nessa lógica que Westin (2003) forneceu uma análise do conflito existente entre a privacidade e o monitoramento presente na sociedade moderna, com o objetivo de promover um entendimento do papel que a privacidade assume diante dos avanços da tecnologia. As pesquisas realizadas pelo autor, no período 1995-2003, dividiram os respondentes em três categorias: fundamentalistas (aqueles que apresentam alto nível de preocupação com relação à proteção de suas informações), pragmáticos (que acreditam nos benefícios de dispor suas informações pessoais às empresas, mas ao mesmo tempo, preocupam-se com a forma como

elas utilizam os seus dados) e despreocupados (baixo nível de preocupação e estão mais dispostos a fornecer suas informações pessoais para as empresas e instituições).

Mai (2016) ressalta a necessidade de inserir a privacidade no contexto do Big Data, em que os indivíduos não só se preocupam com a coleta de dados feita pelas organizações, mas também com a forma como esses dados serão analisados e usados, o que o autor denomina de *datafication*. Com o intuito de propor um modelo de investigação que segue essa lógica do autor, utilizou-se, neste estudo, além da escala de preocupação com a privacidade informacional desenvolvida por Westin (2003), a escala de preocupação com a privacidade na Internet formulada por Castañeda e Montoro (2007), que foi replicada com o intuito de englobar o conhecimento e o controle tanto da coleta de dados, quanto do uso que as empresas fazem desses dados.

Uma vez que o objetivo deste estudo volta-se para o comportamento do consumidor com relação à sua privacidade informacional, buscou-se propor um modelo teórico que avalia a expertise em utilizar a Internet, o risco percebido e o controle informacional como fatores que determinam a relação entre a preocupação com a privacidade e, consequentemente, a confiança e a disposição a fornecer informações on-line. A escolha desses construtos foi feita por abranger o uso da Internet como um todo, não só o comércio eletrônico, mas também o uso de diferentes plataformas on-line. Outros fatores identificados na literatura não foram considerados por basearem-se em uma empresa on-line ou plataforma específica, diferente do objetivo proposto para esta pesquisa.

2.1.1 Expertise em utilizar a Internet

Algumas pesquisas apontam o grau de habilidade dos consumidores de utilizar a Internet como um dos fatores que afetam os diferentes níveis de preocupação com a privacidade (HARRIS; HOYE; LIEVENS, 2003; DINEV; HART, 2005; LOPEZ-NICOLAS; MOLINA-CASTILLO, 2008). Conforme Dinev e Hart (2005), essa expertise inclui: a capacidade do usuário de estabelecer uma conexão on-line; de se orientar de uma maneira autossuficiente em uma página da Internet; de lidar com problemas técnicos; de realizar transações por meio do *e-commerce*; de realizar buscas de maneira rápida e eficiente; de lidar com *spams* ou com vírus de aplicativos danosos; de saber configurar as opções de privacidade e segurança pelo *browser*; e de identificar websites e e-mails enganosos.

Ao utilizar a Internet, os indivíduos ficam vulneráveis a ataques, vírus e invasões, muitas vezes facilitadas pelo baixo grau de conhecimento de uso dos usuários. Diante disso, Harris, Hoye e Lievens (2003), Dinev e Hart (2005) e Lopez-Nicolas e Molina-Castillo (2008) observaram em seus estudos que quanto maior a expertise de utilizar a Internet, menor o nível de preocupação com a privacidade. Isso pode ser explicado pelo fato de que aqueles que apresentam um maior conhecimento técnico são capazes de lidar melhor com os riscos, tendem a customizar seus browsers e aplicativos para protegê-los de vírus e utilizam programas antivírus e contra *malwares*. Assim, aqueles que se autodeclaram com um alto nível de expertise tendem a ter uma menor preocupação com a privacidade informacional.

2.1.1 Risco percebido

Optou-se pela definição proposta por Malhotra, Kim e Agarwal (2004), de que o risco percebido (ou as crenças de risco) consiste na expectativa dos consumidores de que há grandes chances de haver consequências negativas caso decidam fornecer suas informações pessoais às empresas on-line.

A partir desse conceito, alguns estudos consideram o risco percebido como um dos principais fatores que afetam a percepção de privacidade dos indivíduos e a construção da relação de confiança entre os consumidores e as empresas on-line. Ao investigar as percepções de risco e privacidade dos consumidores no comércio eletrônico, Miyazaki e Fernandez (2001)

e Youn (2009) observaram que um maior nível de percepção do risco, com relação ao fornecimento de informações pessoais, gera um maior nível de preocupação com a privacidade, enquanto maiores níveis de percepção dos benefícios ao fornecer informações resultaram em menor nível de preocupação com a privacidade.

2.1.1 Controle informacional

Ao adotarem estratégias de personalização, as empresas podem coletar e utilizar as informações de seus consumidores explicitamente, pedindo para que forneçam os seus dados pessoais, ou implicitamente, compilando ou inferindo as informações do comprador sem o seu consentimento. Nesse sentido, muitas vezes os consumidores só se dão conta de que seus dados foram coletados quando percebem anúncios ou outras formas de comunicação personalizada (TAYLOR; DAVIS; JILLAPALLI, 2009).

No contexto de personalização on-line, Taylor, Davis e Jillapalli (2009) observaram que quando o nível de controle informacional é baixo, maior é o nível de preocupação com a privacidade, o que acaba afetando a confiança e a intenção dos consumidores de adotarem serviços personalizados. Contudo, tendo em vista que este estudo busca examinar o controle informacional de uma forma mais ampla, considerando todas as atividades que são desempenhadas on-line, buscou-se avaliar o desejo dos consumidores pelo controle de suas informações ao invés do controle percebido, uma vez que não será analisado nenhum *website* ou serviço específico. Conforme posto por Phelps, Nowak e Ferrell (2000), quanto maior o desejo pelo controle informacional, maiores são os níveis de preocupação com a privacidade.

2.2 Relação entre a preocupação com a privacidade, a confiança e a disposição a fornecer informações on-line

São diversos os estudos que investigam o dilema do consumidor entre o valor da personalização e a preocupação com a privacidade (SCHOENBACHLER; GORDON, 2002; MALHOTRA; KIM; AGARWAL, 2003; CHELLAPPA; SIN, 2005; AWAD; KRISHNAN, 2006; SHENG; NAH; SIAU, 2008; LIN; LIU, 2012; SUTANTO et al., 2013). Conforme esses autores, a preocupação com a privacidade afeta negativamente a confiança dos consumidores nos serviços tecnológicos on-line e, consequentemente, a disposição a fornecer suas informações pela Internet.

Essa relação se explica pelo fato de que a confiança, no contexto da Internet, pode ser entendida como o grau em que uma pessoa acredita que as suas informações estão protegidas por empresas on-line (MALHOTRA; KIM; AGARWAL, 2004). Tem-se, ainda, a concepção de delegação de confiança, que constitui-se como algo transitivo, de modo que os consumidores delegam sua confiança às empresas, autorizando-as a tomar decisões com base nas suas informações, de modo que essas empresas podem confiar essas informações a terceiros (GRANDISON; SLOMAN, 2000).

Assim, fatores de construção de confiança não só afetam diretamente o uso de serviços de personalização pelos consumidores, como também estão negativamente relacionados com a preocupação com a privacidade. O tipo de informação requisitada pelos *websites* também influencia as percepções e as intenções dos consumidores. Desse modo, quanto maior a preocupação com a privacidade, menor a disposição do consumidor de ser perfilado, seja para serviços ou para anúncios personalizados (MALHOTRA; KIM; AGARWAL, 2004; CHELLAPPA; SIN, 2005; AWAD; KRISHNAN, 2006).

Em suma, através da análise desses estudos, foi possível perceber que a preocupação com a privacidade apresenta uma relação negativa com a confiança, que também afeta negativamente a disposição a fornecer informações, o que foi observado em vários contextos no âmbito on-line.

2.3 Modelo teórico e hipóteses da pesquisa

Com base no levantamento da literatura, o modelo teórico desta investigação propõe a preocupação com a privacidade informacional como um antecedente da relação de confiança e da disposição dos consumidores a fornecer suas informações pessoais on-line. Por sua vez, a preocupação com a privacidade é influenciada pelas características pessoais dos indivíduos: expertise em utilizar a Internet, risco percebido e o desejo pelo controle informacional.

Entende-se ainda que a relação entre as características pessoais e a preocupação com a privacidade pode ser moderada pelos traços de personalidade dos consumidores e pelos seus valores culturais. A personalidade foi mensurada por meio do modelo dos cinco fatores da personalidade, sendo esses: i) extroversão, cujos indivíduos são sociáveis, comunicativos, positivos e ativos; ii) a conscienciosidade, característica daqueles que são mais organizados, responsáveis, diretos e minuciosos; iii) a amabilidade, que são pessoas bondosas, condescendentes, modestas, gentis, altruístas e cooperativas; iv) o neuroticismo, representado por indivíduos que apresentam altos níveis de ansiedade, depressão, raiva e insegurança; e v) a abertura para experiências, traço daqueles que geralmente são mais intelectuais, imaginativos, sensíveis e possuem uma mente aberta (ROCCAS et al., 2002). No que se refere à influência dos valores culturais, neste estudo, o individualismo/coletivismo foi considerado como um fator que modera a relação da expertise em utilizar a Internet, do risco percebido e do controle informacional com a preocupação com a privacidade, uma vez que a natureza dimensional desse construto permite adotá-lo como indicador da cultura nacional (FERREIRA; ASSMAR; SOUTO, 2002).

Além disso, buscou-se avaliar o efeito moderador do comportamento de uso da Internet (intensidade de uso, compras na Internet, cadastro em websites) e do nível de conhecimento sobre o Big Data (percepção de rastreamento para a coleta de dados pelas empresas; percepção de rastreamento para a personalização e o quanto se sabe o que é o Big Data). Essas variáveis foram consideradas como moderadoras, partindo do pressuposto de que esses fatores assumem um nível secundário e influenciam a relação entre as variáveis independentes e dependentes a serem investigadas (MARCONI; LAKATOS, 2003). As hipóteses formuladas estão apresentadas no Quadro 1.

Quadro 1 – Hipóteses da pesquisa

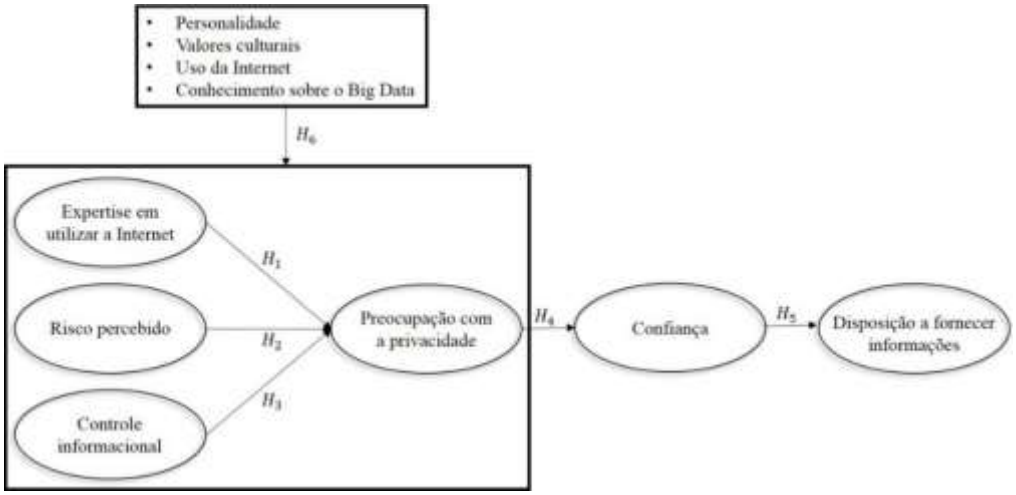
Hipóteses	Autor(es)
H_1: A expertise em utilizar a Internet influencia negativamente a preocupação com a privacidade.	Harris, Hoye e Lievens (2003); Dinev e Hart (2005); Lopez-Nicolas; Molina-Castillo (2008).
H_2: O risco percebido influencia positivamente a preocupação com a privacidade.	Miyazaki e Fernandez (2001); Park e Jun (2003); Malhotra, Kim e Agarwal (2004); Lu, Hsu e Hsu (2005); Youn (2009).
H_3: O controle informacional influencia positivamente a preocupação com a privacidade.	Phelps, Nowak e Ferrell (2000); Malhotra, Kim e Agarwal (2004); Taylor, Davis e Jillapalli (2009).
H_4: A preocupação com a privacidade influencia negativamente a confiança.	Schoenbachler e Gordon (2002); Malhotra, Kim e Agarwal (2004); Chellappa e Sin (2005); Lin e Liu (2012).
H_5: A confiança influencia positivamente a disposição a fornecer informações pessoais.	Schoenbachler e Gordon (2002); Malhotra, Kim e Agarwal (2004); Chellappa e Sin (2005); Lin e Liu (2012).

Fonte: As autoras.

Acrescenta-se ainda a hipótese 6, que analisa a influência das variáveis moderadoras sobre a relação entre as características pessoais (expertise em utilizar a Internet, risco percebido

e controle informacional) e a preocupação com a privacidade. As relações entre os construtos analisados e as hipóteses desta pesquisa estão representadas na Figura 1.

Figura 1 – Modelo teórico



Fonte: As autoras.

3 ASPECTOS METODOLÓGICOS

Este estudo tem caráter quantitativo e descritivo, uma vez que busca descrever e interpretar as características do fenômeno estudado por meio da formulação de hipóteses específicas e do tratamento estatístico dos dados (VIEIRA 2002). O processo de amostragem utilizado foi o não-probabilístico e por conveniência.

O método de coleta de dados foi uma *survey*, que teve como objetivo coletar dados sobre as variáveis moderadoras, os fatores determinantes e os construtos de preocupação com a privacidade, de confiança e de disposição a fornecer informações pessoais on-line, conforme foi proposto no modelo teórico de investigação. Além disso, foram levantadas as características demográficas dos respondentes, o comportamento de uso da Internet (cadastro em *websites* para receber ofertas, compras polo *e-commerce*, uso de mídias sociais e conectividade a aplicativos e plataformas on-line). Por fim, foi colocada uma questão não obrigatória em que foi pedido para que os respondentes definissem Big Data.

Para cada construto levantado no referencial teórico, buscou-se identificar um modelo de instrumento que pudesse ser modificado ou aplicado na íntegra, apresentando um alto grau de confiabilidade, de acordo com o Alfa de Cronbach do instrumento completo. A maioria das escalas foi traduzida para o português e algumas foram adaptadas ao contexto da Internet, mas apenas mudanças mínimas foram necessárias.

As escalas de risco percebido, controle informacional, confiança e disposição a fornecer informações pessoais on-line foram replicadas a partir do estudo de Malhotra, Kim e Agarwal (2004), cujos Alfas de Cronbach (α) originais apresentaram os valores de 0,92, 0,78, 0,78 e 0,83, respectivamente. A escala de expertise, desenvolvida por Ohanian (1990), foi adaptada neste estudo para o contexto da Internet, sendo o α original de 0,82. Foram utilizados os instrumentos completos da personalidade (α original variou entre 0,71 e 0,85), desenvolvido por Passos e Laros (2015) e do individualismo/coletivismo (α = 0,74 para o individualismo e 0,70 para o coletivismo), adotados por Singeliz et al. (1995), Gurhan-Canli e Maheswaran (2000), Agrawal e Maheswaran (2005) e Polyorat e Alden (2005). Com relação à preocupação com a privacidade, foram utilizadas duas escalas: a escala de preocupação com a privacidade na Internet, com α original de 0,80 e 0,89, desenvolvida por Castañeda e Montoro (2007), que foi utilizada na análise estatística do modelo proposto; e a escala de preocupação com a

privacidade informacional, desenvolvida por Westin (2003), com o intuito de obter a categorização da população em fundamentalistas, despreocupados e pragmáticos, com base no nível de concordância com as respostas.

Para a determinação do tamanho da amostra, foram consideradas as condições da análise de dados multivariada, cujas dimensões mínimas, propostas por Hair et al. (2014), apresentam a relação de 10 elementos para cada indicador de cada construto analisado. Nesse sentido, para este estudo, a amostra deveria contemplar, no mínimo, 690 unidades amostrais. Após a verificação dos 871 questionários respondidos, a amostra foi estabelecida com 864 elementos.

Utilizou-se a Modelagem de Equações Estruturais com Mínimos Quadrados Parciais (PLS-SEM), pois além de possibilitar a análise de um conjunto de dados com distribuição não-normal, pretendia-se explicar a variância sobre variáveis dependentes do modelo teórico proposto (HAIR JR. et al., 2005).

Além disso, para analisar a variância dos construtos, considerando as características pessoais dos respondentes, foram realizados testes de comparação de médias (Teste t e ANOVA). Todas as análises foram feitas utilizando-se um intervalo de confiança de 95%, considerado aceitável nos estudos nas Ciências Sociais Aplicadas.

Tendo em vista que o modelo de investigação foi testado utilizando a modelagem de equações estruturais, verificou-se, inicialmente, a colinearidade entre as variáveis por meio da análise da tolerância (TOL) e do fator de inflação da variância (VIF). Para todas as correlações do modelo de investigação, a colinearidade não apresentou problema. Como o grau de colinearidade não foi considerado alto (ou TOL < 0,20 e VIF > 5), não foi necessário excluir ou fazer uma nova combinação dos indicadores. Também não foram encontrados indícios de multicolinearidade para nenhuma das variáveis. Feito o tratamento inicial dos dados, a seguir, é apresentada a análise dos resultados.

4 ANÁLISE DOS RESULTADOS

4.1 Caracterização da amostra

A amostra contemplou um total de 864 respondentes, sendo que 64% é composta pelo público feminino. A idade variou entre 18 e 61 anos, sendo a média de 26 anos, com um desvio padrão de 6,69 anos. A maioria dos respondentes (27%) é proveniente do estado de Minas Gerais, seguido de 17% do estado de São Paulo e 10% do Rio Grande do Sul. Contudo, a amostra abrangeu respondentes de todas as regiões brasileiras, sendo a maioria da Região Sudeste (46%), seguida pela Região Sul (23%), Nordeste (16%), Centro-Oeste (12%) e Norte (3%). Além disso, 66% afirmaram morar no interior de seus respectivos estados, enquanto 34% moram na capital.

Quanto ao estado civil, a maioria (82%) declarou-se solteira. A análise segue com 16% casados ou em uma união estável, 2% separados, divorciados ou desquitados e 0,2% são viúvos. A maioria (52%) também afirmou ter Segundo Grau completo/Ensino Superior incompleto. Apenas 0,6% dos respondentes apresentam um nível de escolaridade inferior ao segundo grau completo. A média da renda familiar mensal dos respondentes correspondeu a R$ 3.667,58.

Além da caracterização do perfil demográfico, buscou-se analisar as experiências e o comportamento dos respondentes no âmbito on-line. A maioria (69%) declarou que gasta de 2 a 6 horas por dia na Internet e 95% afirmaram que já realizaram compras na Internet e pretendem continuar comprando. Além disso, observou-se que os 4% que afirmaram nunca ter realizado compras no comércio eletrônico, pensam em comprar no futuro. Também foi questionado aos respondentes se eles já se cadastraram em *websites* para receber ofertas ou conteúdo personalizado, sendo que a resposta foi afirmativa para 91% dos respondentes, compatível com o número de respondentes que possuem a prática de realizar compras na Internet.

Com relação à percepção dos respondentes quanto ao rastreamento digital, 96% acreditam que as atividades que realizam na Internet são rastreadas por empresas e 98% acreditam que as informações que fornecem on-line são utilizadas pelas empresas para a indicação de produtos ou ofertas. Embora a percepção de rastreamento digital tenha sido alta, a maioria (67%) não tem conhecimento do que é o Big Data.

Foram obtidas 303 respostas ao se perguntar qual seria a definição de Big Data aos respondentes. Desse número, 88% responderam acreditar que Big Data está relacionado a uma grande quantidade de dados. Uma minoria (7,5%) definiu Big Data como uma grande quantidade de dados que exigem técnicas especializadas para gerar informações, mostrando maior conhecimento do termo, e 12% foram bastante distintas do significado correto, dentre as quais, a mais incidente foi que Big Data significa uma data importante, a data de um grande evento ou um dia que as empresas on-line realizam promoções. Outras respostas foram o "dia em que acontecerá uma sobrecarga da Internet ou o '*boom'* da Internet", "um aplicativo de *e-commerce*" e "algo que protege as informações dos usuários on-line".

4.2 Atitudes com relação à preocupação com a privacidade

Ao desenvolver o construto de preocupação com a privacidade informacional, Westin (2003) identificou as três diferentes atitudes (fundamentalistas, pragmáticos e despreocupados) analisando o número de respondentes por nota de concordância de cada indicador (KING, 2014). Neste estudo, alguns itens da escala original foram adaptados, considerando o cenário on-line. A partir daí, aqueles que indicaram um baixo grau de concordância (de 1 a 3 pontos) foram considerados despreocupados. Aqueles que avaliaram cada item com alto grau de concordância (de 5 a 7 pontos) foram classificados como fundamentalistas. E aqueles que indicaram "não concordo, nem discordo" (4 pontos) foram considerados pragmáticos. Com base nos dados da pesquisa, a maioria dos respondentes indicou um alto nível de preocupação com a privacidade (67%), caracterizando-se como fundamentalistas, enquanto a minoria da amostra (12%) apresentou um baixo nível de preocupação com a privacidade (despreocupados).

Estes resultados são diferentes daqueles encontrados por Westin (2003), que identificou a prevalência dos pragmáticos, podendo ser explicada não só pelo contexto tecnológico no período em que o autor realizou a pesquisa, mas também pelas características culturais dos Estados Unidos. Porém, uma análise longitudinal de seus estudos já apontava para uma migração acentuada dos despreocupados para os fundamentalistas – de 25 para 34% – e em menor escala para pragmáticos – de 55 para 58%, no período de 1995 a 2002.

Uma outra possível explicação para o grande número de respondentes fundamentalistas corresponde às mudanças recentes no cenário virtual, em que o surgimento das mídias sociais no final da metade dos anos 2000 fizeram com que os indivíduos passassem a lidar com uma maior disponibilização e exposição de suas informações pessoais, de modo que esses dados são constantemente coletados e utilizados pelas empresas para a oferta de produtos e serviços personalizados. Isso veio acompanhado de um menor controle informacional, tendo em vista que essas empresas utilizam-se não somente de informações que são cedidas diretamente pelos consumidores, mas também dos dados que são gerados por meio de cookies ou do histórico de navegação dos usuários.

4.3 Teste das hipóteses da pesquisa

4.3.1 Análise Fatorial Exploratória (AFE)

Para testar as hipóteses do modelo de investigação proposto nesta pesquisa, inicialmente, fez-se necessário validar as escalas utilizadas para a coleta de dados. Essa validação se deu em três fases: análise da Medida de Adequação da Amostra (MSA) e do teste de esfericidade de Bartlett; análise das comunalidades e da matriz de rotação ortogonal; e a análise da confiabilidade das escalas.

Os construtos que compõem o modelo (expertise em utilizar a Internet – EXP, risco percebido – RIS, controle informacional – CTL, preocupação com a privacidade na Internet – PPI, confiança – COM, e disposição a fornecer informações pessoais on-line – DFI) apresentaram um valor MSA = 0,859 e o teste de esfericidade de Bartlett apresentou $p < 0,001$, confirmando a adequação e a suscetibilidade dos dados para a Análise Fatorial.

Em uma segunda etapa, analisou-se as comunalidades dos fatores. Além disso, para identificar a estrutura latente das relações, optou-se pelo método de rotação ortogonal VARIMAX. Todos os fatores (indicadores das variáveis exógenas e endógenas do modelo proposto) apresentaram comunalidades maiores do que 0,50, indicando, dessa forma, cargas fatoriais significantes. Além disso, a AFE indicou 7 componentes com autovalores acima de 1, que explicam 68,94% da variância do fenômeno.

Também observou-se que todos as cargas elevadas foram combinadas em um único componente não apresentando cargas cruzadas. Contudo, variáveis PPI5, PPI6, PPI7 e PPI8 foram combinadas no Componente 4, enquanto as variáveis PPI1, PPI2, PPI3 e PPI4 foram combinadas no Componente 6. Quando isso ocorre, Hair Jr et al. (2005) sugere a designação de diferentes nomes para os componentes. Uma interpretação substancial da escala PPI permite perceber que PPI1, PPI2, PPI3 e PPI4 se referem à preocupação dos respondentes com relação ao controle de suas informações na Internet, enquanto PPI5, PPI6, PPI7 e PPI8 corresponde à preocupação com o uso que os websites que coletam os seus dados fazem deles. Desse modo, a nomeação definida para o Componente 4 foi "preocupação com o controle das informações pessoais na Internet" (PPI_Controle) e para o Componente 6 foi "preocupação com o uso das informações pessoais pelos *websites*" (PPI_Uso).

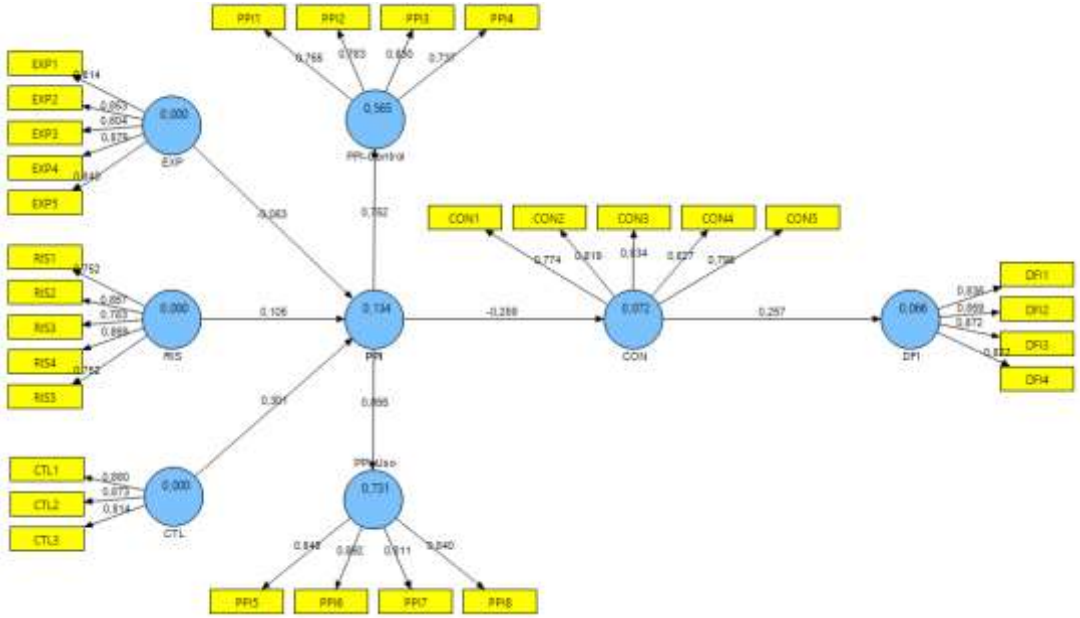
Por fim, buscou-se verificar a confiabilidade e a validação das escalas que compõem o modelo de investigação proposto. As escalas analisadas apontaram um alto grau de confiabilidade (α_{EXP} = 0,90; α_{RIS} = 0,86; α_{CTL} = 0,82; ; α_{PPI} = 0,80; ; $\alpha_{PPI_Controle}$ = 0,78; ; α_{PPI_uso} = 0,87; ; α_{CON} = 0,87; ; α_{DFI} = 0,87).

Ao realizar a Análise Fatorial Exploratória dos dados, a adequação da amostra, os níveis de intercorrelação entre as variáveis, os valores das comunalidades, a estrutura latente das relações e a confiabilidade das escalas apresentaram valores consistentes para dar prosseguimento à Modelagem de Equações Estruturais (SEM), apresentada na sequência.

4.3.2 Modelagem de Equações Estrutuais (PLS-SEM)

O modelo de mensuração foi elaborado a partir do modelo de investigação proposto nesta pesquisa, em que foram avaliadas as validades convergente e discriminante dos dados. Tendo em vista que a matriz de rotação ortogonal gerou dois componentes de preocupação com a privacidade (PPI_Controle e PPI_Uso), PPI foi colocado como um construto de segunda ordem. Feito isso, para a avaliação da validade convergente, foram verificadas as cargas fatoriais dos indicadores de cada construto mensurado. Os valores dos indicadores variaram entre 0,74 e 0,89, não havendo necessidade de excluir nenhum item. Esses valores estão apresentados na Figura 1.

Figura 1 – Modelo de mensuração

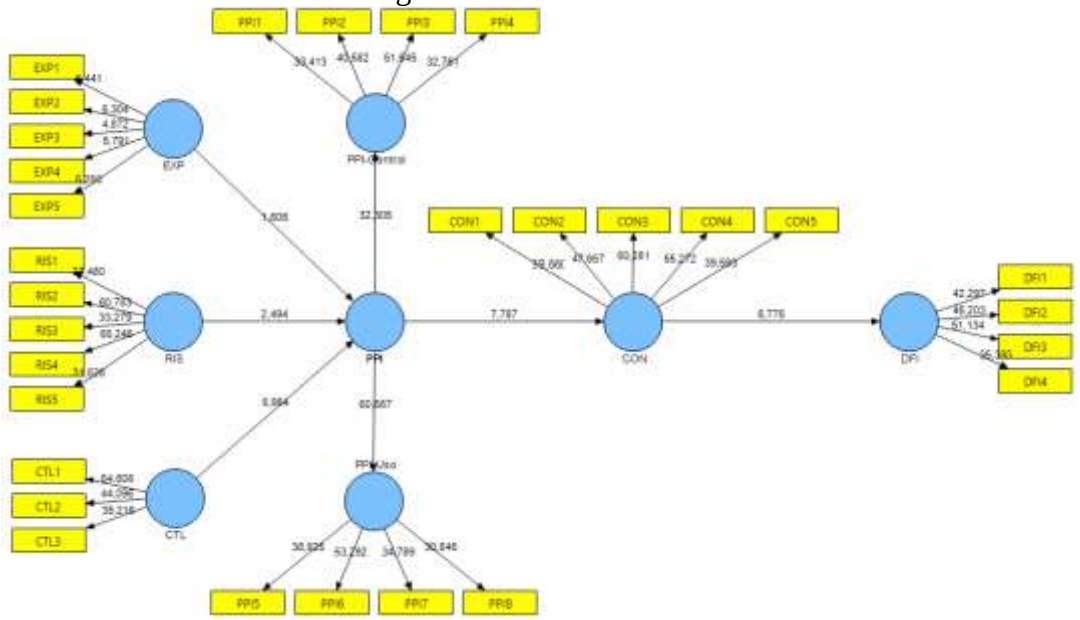


Fonte: Dados da pesquisa.

Em seguida, ao verificar os valores da confiabilidade composta (ρ_c) de cada construto, observou-se que todos apresentaram ρ_c> 0,708, sendo valores aceitáveis, conforme Hair et al. (2014). Prosseguindo-se à validade convergente, buscou-se verificar os valores da Variância Média Extraída (AVE). Exceto para PPI, todos os outros construtos apresentaram valores AVE > 0,60. O valor da AVE de PPI foi de 0,43, mas optou-se manter a Preocupação com a Privacidade Informacional com base na orientação de Bollen (1989), que admite valores para AVE > 0,40. A validade discriminante também foi confirmada, uma vez que não foram identificadas cargas cruzadas.

Tendo em vista que o modelo de mensuração foi satisfatório, buscou-se examinar o modelo estrutural com o intuito de identificar as relações causais existentes entre as variáveis. Na Figura 2 é apresentado o modelo estrutural gerado.

Figura 2 – Modelo estrutural

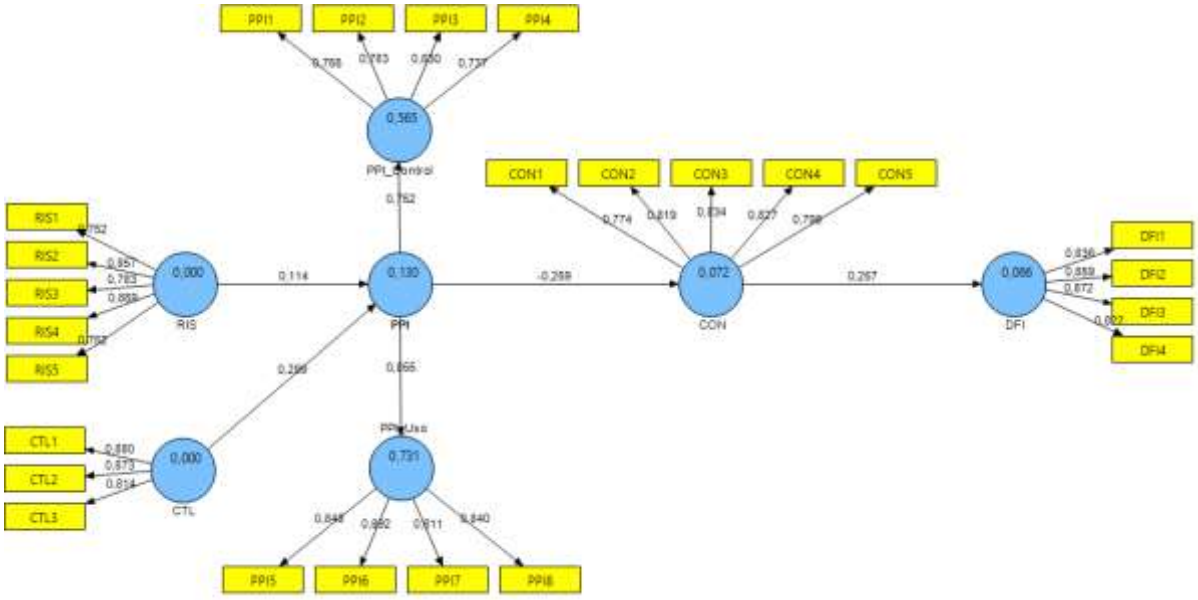


Fonte: Dados da pesquisa.

Para avaliar as relações entre as variáveis, foi realizada uma análise do coeficiente de caminho, considerando como significantes valores maiores que 0,10, e do *t value* (> 1,96), obtido utilizando a técnica de *bootstrap* com o mesmo número de casos da amostra. Com base nos valores apresentados no modelo estrutural, apenas para a relação EXP → PPI, o coeficiente de caminho foi menor que 0,10 e $t < 1{,}96$, rejeitando a hipótese H_1. Uma vez que a relação EXP → PPI não foi significante, EXP foi excluída do modelo e, novamente, foi gerado o modelo de mensuração para avaliar os coeficientes de determinação (R^2), apresentado na Figura 3.

Os valores R^2 das variáveis endógenas após a exclusão da expertise em utilizar a Internet (EXP), com base nas colocações de Cohen (1988), não apresentaram um efeito grande. Em complementaridade, foram avaliados os coeficientes de redundância cross-validado (Q^2) e do tamanho do efeito (f^2). Embora Q^2 tenha apresentado valores adequados (>0,00) para todas as variáveis endógenas, o coeficiente do tamanho do efeito f^2 foi pequeno (próximos de 0,02), ou seja, a exclusão da variável EXP apresentou um efeito pequeno sobre o modelo. A análise prosseguiu para a avaliação do efeito das variáveis moderadoras, partindo do pressuposto de que esse efeito pode implicar em maiores coeficientes (R^2, Q^2 e f^2) das relações entre as variáveis exógenas e endógenas. Contudo, o modelo não apresentou efeito de nenhuma variável moderadora.

Figura 3 – Modelo de mensuração após exclusão da variável EXP



Fonte: Dados da pesquisa.

Observa-se, de um modo geral, que o modelo de investigação proposto, apesar de não sofrer efeito de nenhuma das variáveis moderadoras e de ter rejeitado a hipótese 1 (ou seja, a expertise em utilizar a Internet não afeta a preocupação com a privacidade), mostrou-se adequado para a acurácia preditiva. O risco percebido e o controle informacional são fatores que determinam a preocupação com a privacidade, que afeta negativamente a confiança dos indivíduos, e a confiança influencia positivamente a disposição a fornecer informações pessoais na Internet.

Outros modelos alternativos foram testados com base na literatura levantada. No entanto, para ambos, o R^2 das variáveis endógenas apresentaram valores menores do que os valores do modelo originalmente proposto e, desse modo, não foi considerado parcimonioso. Portanto, para este estudo, o modelo teórico proposto mostrou ser o mais adequado.

5. DISCUSSÃO DOS RESULTADOS

De um modo geral, o modelo teórico proposto neste estudo mostrou estar em conformidade com a literatura levantada. Exceto para o construto expertise em utilizar a Internet (EXP), o risco percebido (RIS) e o controle informacional (CTL) são fatores que determinam a preocupação com a privacidade (PPI). A preocupação com a privacidade (PPI), por sua vez, afeta negativamente a confiança (CON), ou seja, quanto maior for o nível de preocupação com a privacidade, menos os indivíduos acreditam que as empresas on-line fornecem proteção e segurança ao coletarem os seus dados (SCHOENBACHLER; GORDON, 2002; MALHOTRA; KIM; AGARWAL, 2004; CHELLAPPA; SIN, 2005; LIN; LIU, 2012). E, por fim, a confiança influencia a disposição a fornecer informações pessoais (DFI), de modo que, quanto maior for a confiança, maior a disposição dos consumidores de abrirem mão de suas informações pessoais (SCHOENBACHLER; GORDON, 2002; MALHOTRA; KIM; AGARWAL, 2004; CHELLAPPA; SIN, 2005; LIN; LIU, 2012).

Não foram observados efeitos moderadores de outras variáveis, que também pode ser explicado pela baixa variabilidade das respostas dadas pelos respondentes. Por meio da análise descritiva dos construtos, observou-se que a maioria dos participantes apresentam altos níveis de percepção de risco (78,4%), de desejo pelo controle informacional (93,3%) e, consequentemente, de preocupação com a privacidade (95,7%). Como resultado, apresentaram baixos níveis de confiança (apenas 17,9% apresentaram níveis de confiança acima do ponto médio da escala) e disposição a fornecer informações on-line (9,5%).

Isso ficou evidenciado ao analisar as atitudes com relação à preocupação com a privacidade propostos por Westin (2003), em que a maioria dos respondentes foi classificada como fundamentalista. No entanto, ao observar o comportamento de uso da Internet, notou-se, contraditoriamente, que mais de 70% da amostra: já realizaram compras através de comércio eletrônico e pretendem continuar comprando; já se cadastraram em *websites* para receber ofertas ou conteúdo personalizado; fazem uso de mídias sociais, *e-mails*, *messengers*, plataformas de armazenamento em nuvem e de aplicativos (como Internet *banking* e aplicativos de geolocalização); e têm consciência do rastreamento digital feito pelas empresas.

Essa contradição pode ser entendida pelo fato de que, apesar dos respondentes acreditarem que fornecer suas informações pessoais pela Internet envolve riscos, eles enxergam os benefícios da personalização. Além disso, conforme posto por Phelps, Nowak e Ferrell (2000) e Schoenbachler e Gordon (2002), a disposição a fornecer informações pessoais on-line depende de uma série de fatores, como o tipo de dados que são fornecidos. Os consumidores, geralmente, apresentam baixa resistência a fornecer dados demográficos ou de hábitos de consumo do que dados financeiros, ou, até mesmo, de suas características individuais.

Neste estudo, observou-se, por exemplo, que as mulheres e aqueles com menor nível de escolaridade apresentaram uma maior percepção de risco e um maior desejo pelo controle de suas informações pessoais, apesar de não serem identificadas relações significativas entre gênero e escolaridade, e a preocupação com a privacidade, com a confiança e com a disposição a fornecer informações pessoais. Além disso, apesar da personalidade não apresentar efeito moderador no modelo proposto, ao fazer análise de correlação, foi possível observar que a amabilidade e a abertura para experiências se relacionam positivamente com a preocupação com a privacidade, enquanto a extroversão apresentou uma relação positiva com a disposição a fornecer informações on-line ($p < 0,001$).

No que se refere aos valores culturais, observou-se que tanto o individualismo, quanto o coletivismo se associaram positivamente com a preocupação com a privacidade. Conforme Lowry, Cao e Everard (2011), os indivíduos de sociedades mais individualistas apresentam menores níveis de preocupação com a privacidade e maiores níveis de disposição a fornecer informações pessoais, enquanto os indivíduos coletivistas tendem a ser mais preocupados com a proteção de suas informações e, portanto, menos dispostos a fornecê-las. No entanto, os

autores destacam que é comum esses resultados serem divergentes. Sobretudo, notou-se que o coletivismo apresentou uma relação mais forte com a preocupação com a privacidade do que o individualismo.

Vale considerar também que apesar da expertise em utilizar a Internet ter sido excluída do modelo, observou-se que a variável afeta negativamente a preocupação com a privacidade, conforme já era esperado. Também foi possível perceber uma relação significativa com a disposição a fornecer informações pessoais. Isso porque os indivíduos que apresentam um maior nível de expertise são mais capazes de lidar com ameaças, como vírus, *spams* e *malwares* e, geralmente, configuram as opções de privacidade e segurança pelo *browser* (DINEV; HART, 2005).

Com base nessas considerações, o modelo permitiu verificar que, no atual cenário virtual, em que estão sendo desenvolvidas ferramentas cada vez mais sofisticadas para a coleta e interpretação do Big Data, os respondentes mostraram-se altamente preocupados com a privacidade, dado que acreditam que fornecer suas informações pessoais on-line envolve riscos e sentem uma maior necessidade de um controle informacional. Isso acaba afetando a confiança com relação aos *websites* que coletam os seus dados e, consequentemente, promove uma maior resistência a fornecer informações pessoais.

Observou-se que, apesar de grande parte dos respondentes não saber o que é o Big Data, mais de 90% percebem que seus dados são rastreados por empresas on-line para a divulgação de ofertas e serviços personalizados que, através dos resultados, apresentou uma relação significativa com a confiança, sendo que aqueles que acreditam que seus dados são rastreados por empresas e que são utilizados para a personalização on-line, mostraram um menor grau de confiança nos *websites*. Desse modo, este estudo apresenta importantes contribuições teóricas e gerenciais que serão apresentadas na próxima seção.

6. CONSIDERAÇÕES FINAIS

Com base na análise e discussão dos resultados, esta pesquisa apresenta contribuições teóricas ao propor um modelo que explica como o risco percebido e a necessidade pelo controle informacional afetam a preocupação com a privacidade, a confiança e, consequentemente, a disposição dos consumidores a fornecer informações pela Internet. Muitos estudos no Brasil, que analisam o contexto virtual, voltam-se para fatores que impactam a atitude e a intenção de compra em comércio eletrônico ou de uso de plataformas on-line, sem levar em conta o *trade-off* que os indivíduos fazem entre a privacidade e o valor dos serviços personalizados.

Nesse sentido, este estudo também apresenta importantes implicações gerenciais. Geralmente, as empresas oferecem serviços personalizados apenas para aqueles consumidores que frequentemente acessam seus *websites* ou que pesquisam pelos seus produtos. Somente por meio dessas informações não é possível avaliar a intenção de compra ou de uso de suas plataformas on-line, nem a intenção dos consumidores de fornecer seus dados pessoais. Além disso, alguns estudos enfatizam a importância de analisar os fatores que afetam a preocupação com a privacidade para a construção de uma relação de confiança e para que a atitude dos consumidores com relação às empresas on-line seja positiva (SCHOENBACHLER; GORDON, 2002; MALHOTRA; KIM; AGARWAL, 2004; CHELLAPPA; SIN, 2005; TAYLOR; DAVIS; JILLAPALLI, 2009; LIN; LIU, 2012).

Por fim, uma vez que os resultados indicaram que uma grande parte da amostra apresenta um alto nível de preocupação com a privacidade e, desse modo, um baixo nível de confiança e disposição a fornecer informações pessoais, os gestores devem adotar estratégias que reduzam as percepções de risco, como melhorar as políticas de privacidade, aumentar a transparência, obter a certificação de uma terceira parte e garantir que os consumidores tenham acesso e controle de suas informações pessoais que são coletadas, uma vez que observou-se também um alto nível de desejo pelo controle informacional.

As implicações teóricas e gerenciais deste estudo podem ser ponderadas pelas suas limitações. Apesar da amostra abarcar respondentes com diferentes características demográficas, concentrou-se, em grande parte, em estudantes universitários, dado o modo em que os dados foram coletados (por meio da divulgação da pesquisa em vários grupos de universidades e por meio da técnica de amostragem bola de neve, que capta respondentes que apresentam, de um modo geral, as mesmas características).

Destaca-se também outros construtos levantados na literatura que não foram explorados nesta pesquisa, tendo em vista que buscou-se entender o comportamento de utilização da Internet como um todo e não de um *website* ou plataforma específica, inviabilizando a análise de outros fatores que afetam a preocupação com a privacidade dos consumidores. Assim, para futuras pesquisas, sugere-se a identificação de outros fatores que afetam a preocupação com a privacidade, tais como sensibilidade à informação, consciência das políticas de privacidade, familiaridade com a empresa on-line, intenção de uso/compra e a reputação.

REFERÊNCIAS

ADOMAVICIUS, G.; TUZHILIN, A. Personalization technologies: a process-oriented perspective. **Communications of the ACM**, v. 48, n. 10, p. 83-90, 2005.

AGRAWAL, N.; MAHESWARAN, D. The effects of self-construal and commitment on persuasion. **Journal of Consumer Research**, v. 31, p. 841-849, mar. 2005.

ALOYSIUS, J. A.; HOEHLE, H.; VENKATESH, V. Exploiting big data for customer and retailer benefits: a study of emerging mobile checkout scenarios. **International Journal of Operations and Production Management**, v. 36, n. 4, p. 467-486, 2016.

AWAD, N. F.; KRISHNAN, M. S. The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. **MIS Quarterly**, v. 30, n. 1, p. 13-28, 2006.

BOLLEN, K. A. **Structural Equations with Latent Variables**. New York: Wiley-Interscience publication, 1989.

BOYD, D.; CRAWFORD, K. Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, v. 15, n. 5, p. 662-979, 2012.

CASTAÑEDA, J. A.; MONTORO, F. J. The effect of Internet general privacy concern on customer behavior. **Electronic Commerce Research**, v. 7, n. 2, p. 117-141, 2007.

CHELLAPPA, R. K.; SIN, R. G. Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. **Information Technology and Management**, v. 6, pp. 181-202, 2005.

CHEN, C. L. P.; ZHANG, C. Y. Data-intensive applications, challenges, techniques and technologies: a survey on Big Data. **Information Sciences**, v. 275, p. 314-347, 2014.

CHEN; H.; CHIANG, R. H.; STOREY, V. C. Business intelligence and analytics: from Big Data to Big Impact. **MIS Quarterly**, v. 36, n. 4, p. 1165-1188, 2012.

COHEN, J. **Statistical power analysis for behavioral sciences.** 2 ed. New Jersey: Lawrence Erbaum Associates, 1988.

COULDRY, N.; TUROW, J. Advertising, big data and the clearance of the public realm: marketers' new approaches to the content subsidy. **International Journal of Communication**, v. 8, p. 1710-1726, 2014.

DINEV, T.; HART, P. Internet privacy concerns and social awareness as determinants of intention to transact. **International Journal of Electronic Commerce**, v. 10, n. 2, p. 7-29, 2005.

EASTIN, M. S.; BRINSON, N. H.; DOOREY, A.; WILCOX, G. Living in a big data world: predicting mobile commerce activity through privacy concerns. **Computers in Human Behavior**, v. 58, p. 214-220, 2016.
FERREIRA, M. C.; ASSMAR, E. M. L.; SOUTO, S. O. O individualismo e o coletivismo como indicadores de culturas nacionais: convergências e divergências teórico-metodológicas. **Psicologia em Estudo**, v. 7, n. 1, p. 81-89, 2002.
FLORIDI, L. The ontological interpretation of informational privacy. **Ethics and Information Technology**, v. 7, p. 185-200, 2005.
GRANDISON, T.; SLOMAN, M. A survey of trust in Internet applications. **IEEE Communications Surveys & Tutorials**, v. 3, n. 4, p. 2-16, 2000.
GURHAN-CANLI, Z.; MAHESWARAN, D. Cultural variations in country of origin effects. **Journal of Marketing Research**, v. 37, n. 3, p. 309-317, 2000.
HAIR JR, J. F.; ANDERSON, R. E.; TATHAM, R. L. BLACK, W. C. **Análise multivariada de dados**. 5 ed. Porto Alegre: Bookman, 2005.
HAIR JR, J. F.; HULT, G. T. M.; RINGLE, C. M.; SARSTEDT, M. **A primer on partial least squares structural equation modeling (PLS-SEM)**. Thousand Oaks: SAGE Publications, 2014.
HARRIS, M. M.; HOYE, G. van; LIEVENS, F. Privacy and attitudes towards Internet-based selection systems: a cross-cultural comparison. **International Journal of Selection and Assessment**, v. 1, n. 2-3, p. 230-236, 2003.
KASPER, D. V. S. The Evolution (or Devolution) of Privacy. **Sociological Forum**, v. 20, n. 1, pp. 69-92, mar. 2005.
KING, J. Taken out of context: and empirical analysis of Westin's Privacy Scale. In: Symposium on Usable Privacy and Security (SOUPS). **Anais...**, 2014.
LI, H.; SARATHY, R.; XU, H. The role of affect and cognition on online consumers' decision to disclose personal information to unfamiliar online vendors. **Decision Support Systems**, v. 51, n. 3, p. 434-445, 2011.
LIN, S.W.; LIU, Y. C. The effects of motivations, trust, and privacy concern in social networking. **Service Business**, v. 6, n. 4, p. 411-424, 2012.
LOPEZ-NICOLAS, C.; MOLINA-CASTILLO, F. J. Customer knowledge management and e-commerce: the role of customer perceived risk. **International Journal of Information Management**, v. 28, n. 2, p. 102-113, 2008.
LOWRY, P. B; CAO, J.; EVERARD, A. Privacy concerns versus desire for interpersonal awareness in driving the use of self-disclosure technologies: the case of instant messaging in two cultures. **Journal of Management Information Systems**, v. 27, n. 4, p. 163-200, 2011.
LU, H. P.; HSU, C. L.; HSU, H. Y. An empirical study of the effect of perceived risk upon intention to use online applications. **Information Management & Computer Security**, v. 13, n. 2, p. 106-120, 2005.
MALHOTRA, N. K.; KIM, S. S.; AGARWAL, J. Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. **Information Systems Research**, v. 15, n. 4, p. 336-355, 2004.
MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 5 ed. São Paulo: Atlas, 2003.
MARGULIS, S. T. Privacy as a Social Issue and Behavioral Concept. **Journal of Social Issues**, v. 59, n. 2, pp. 243-261, 2003.
MICHAEL, K.; MILLER, Keith W. Big Data: New Opportunities and New Challenges. **IEEE Computer Society**, 2013.
MIYAZAKI, A. D.; FERNANDEZ, A. Consumer perceptions of privacy and security risks for online shopping. **The Journal of Consumer Affairs**, v. 35, n. 1, p. 27-44, 2001.
MOORE, A. Defining Privacy. **Journal of Social Philosophy**, v. 39, n. 3, pp. 411-428, 2008.

MYERSCOUGH, S.; LOWE, B.; ALPERT, F. Willingness to provide personal information online: the role of perceived privacy risk, privacy statements and brand strength. **Journal of Website Promotion**, v. 2, n. 1-2, p. 115-140, 2006.
OHANIAN, R. Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. **Journal of Advertising**, v. 19, n. 3, p. 39-52, 1990.
OSATUYI, B. Personality traits and information privacy concern on social media platforms. **Journal of Computer Information Systems**, v. 55, n. 4, p. 11-19, 2015.
PARK, C.; JUN, J. K. A cross-cultural comparison of Internet buying behavior: effects of Internet usage, perceived risks, and innovativeness. **International Marketing Review**, v. 20, n. 5, p. 534-553, 2003.
PASSOS, M. F. D.; LAROS, J. A. Construção de uma escala reduzida de cinco grandes fatores da personalidade. **Avaliação Psicológica**, v. 14, n. 1, p. 115-123, 2015.
PENTLAND, A. The Data-Driven Society. **Scientific American**, v. 309, n. 4, p. 78-83, 2013.
PHELPS, J.; NOWAK, G.; FERRELL, E. Privacy concerns and consumer willingness to provide personal information. **Journal of Public Policy & Marketing**, v. 19, n. 1, p. 27-41, 2000.
POLYORAT, K.; ALDEN, D. Self-construal and need-for-cognition advertising in Thailand and the United States. **Journal of Advertising**, v. 34, n. 1, p. 37-48, 2005.
ROCCAS, S.; SAGIV, L.; SCHWARTZ, S. H.; KNAFO, A. The big five personality factors and personal values. **Personality and Social Psychology Bulletin**, v. 28, n. 6, p. 789-801, 2002.
SCHOENBACHLER, D. D.; GORDON, G. L. Trust and customer willingness to provide information in database-driven relationship marketing. **Journal of Interactive Marketing**, v. 16, n. 3, p. 2-16, 2002.
SHENG, H.; NAH, F. F. H.; SIAU, K. An experimental study on ubiquitous commerce adoption: impact of personalization and privacy concerns. **Journal of the Association for Information Systems**, v. 9, n. 6, p. 344-376, 2008.
SOLOVE, D. J. **Understanding Privacy**. Harvard University Press: Cambridge, MA, 2008.
SUTANTO, J.; PALME, E.; TAN, C. H.; PHANG, C. W. Adressing the personalization-privacy paradox: an empirical assessment from a field experiment on smartphone users. **MIS Quarterly**, v. 37, n. 4, p. 1141-1164, 2013.
TAYLOR, D. G.; DAVIS, D. F.; JILLAPALLI, R. Privacy concern and online personalization: the moderating effects of information control and compensation. **Electronic Commerce Resarch**, v. 9, n. 3, p. 203-223, 2009.
VIEIRA, V. A. As tipologias, variações e características da pesquisa de marketing. **Revista FAE**, v. 5, n. 1, p. 61-70, 2002.
WESTIN, A. F. Social and Political Dimensions of Privacy. **Journal of Social Issues**, v. 59, n. 2, pp. 431-453, 2003.
YOUN, S. Determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents. **Fall**, v. 43, n. 3, p. 389-418, 2009.
ZIMMER, M. ‘‘But the data is already public’’: on the ethics of research in Facebook. **Ethics Inf Technol,** v. 12, p.313-325, 2010.
ZWITTER, A. Big Data ethics. **Big Data & Society**, 2014.