

Avaliação da Utilização das Escalas SERVQUAL e SERVPERF no Brasil: uma análise da adequação do uso de escalas existentes em pesquisas empíricas.

MARCUS RODRIGUES PEIXOTO

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO (PUC-RIO)

ANGELA MARIA CAVALCANTI DA ROCHA

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO (PUC-RIO)

LAÍS RODRIGUES DE OLIVEIRA

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO (PUC-RIO)

CAROLINA TERRA QUIRINO DA COSTA

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO (PUC-RIO)

Avaliação da Utilização das Escalas SERVQUAL e SERVPERF no Brasil: uma análise da adequação do uso de escalas existentes em pesquisas empíricas.

INTRODUÇÃO

O setor de serviços, caracterizado por sua heterogeneidade, potencialidade e dinamismo, tem elevado, no decorrer das últimas décadas, sua participação no Produto Interno Bruto e no emprego da economia brasileira (Alves, Madeira, & Macambira, 2012). Para Cardoso e Perobelli (2013, p. 2), “observando a evolução e crescimento do setor de serviços no Brasil, podemos apontar para um movimento no sentido de uma ‘economia de serviços’, seguindo o padrão mundial”. O setor é responsável por fornecer insumos para a inovação na produção, sendo “impossível pensar o desenvolvimento econômico do Brasil sem um setor de serviços eficiente e capaz de ampliar as condições para que a inovação seja realizada na economia” (IPEA, 2006). A receita operacional líquida do setor de serviços somou 1,5 trilhão de reais em 2017 (IBGE, 2017).

Na esteira do crescimento do setor, a associação do aumento da lucratividade com a gestão de serviços, somada a forte competitividade da área, geraram a necessidade de maior foco na gestão da qualidade. A partir da compreensão de aspectos que apontavam a subjetividade e a intangibilidade da percepção da qualidade dos serviços pelos clientes, como satisfação e lealdade (Zeithmal, Parasuraman, & Berry, 1990), gerou-se a necessidade de melhor compreender as expectativas e a apreensão da qualidade do serviço do ponto de vista dos clientes.

Em paralelo a esse movimento, a utilização de escalas para a mensuração de construtos não observáveis vinha se tornando cada vez mais difundida nas ciências sociais e da saúde, passando a ser prática comum na mensuração de construtos em marketing (Netemeyer, Bearden, & Sharma, 2003). A praticidade e a eventual segurança de sua utilização, porém, pode ser comprometida por possível inadequação para refletir o construto de interesse do pesquisador ou por níveis inaceitáveis de alteração em sua estrutura e conteúdo. Em tais casos, o risco da geração de achados de pesquisa equivocados é significativo. Nesta linha, identificou-se uma lacuna nos estudos teóricos sobre o uso de escalas no Brasil, que tendem ou a realizar um estado da arte sobre as principais escalas utilizadas em determinada área do conhecimento e nicho de estudo, ou a analisar o desenvolvimento de uma nova escala para medir determinado construto. Não é estudado, entretanto, como escalas previamente existentes têm sido adaptadas e utilizadas pelos pesquisadores. Com o propósito de guiar os pesquisadores que utilizam escalas já existentes, decidiu-se realizar um estudo sobre o uso de duas escalas que medem percepção de qualidade de serviços.

O construto qualidade percebida em serviços é um dos mais pesquisados em marketing. Diversas escalas foram desenvolvidas para medir o construto (Abreu, Antonialli, & Almeida, 2019). Dentre elas, destacam-se as escalas SERVQUAL e SERVPERF, utilizadas amplamente, inclusive em pesquisas no Brasil. Apesar de seu amplo uso, nota-se uma ausência de avaliações sobre a forma pela qual tais escalas vêm sendo utilizadas, gerando potenciais riscos para pesquisadores e gestores na aplicação e utilização dos achados destes estudos empíricos. A adaptação das escalas ao setor de serviços específico ou à cultura local é desejável (Sperber, 2004). Existem, entretanto, protocolos e processos de avaliação e validação que devem ser respeitados (Churchill, 1979; Netemeyer *et al.*, 2003).

Sendo assim, este trabalho tem por objetivo avaliar a adequação do uso das escalas SERVQUAL e SERVPERF para mensuração do construto qualidade percebida em serviços em artigos científicos brasileiros de natureza descritiva e quantitativa. Apesar de a mensuração da qualidade percebida ser um tema ainda em evolução, incluindo a geração de diversos novos instrumentos de mensuração nos últimos anos, a escolha das escalas SERVQUAL e SERVPERF se deve a sua ampla utilização, fator primordial para garantir uma maior solidez ao

protocolo que será elaborado para avaliar a aplicação de escalas já existentes em pesquisas no Brasil.

Entende-se que, além de proporcionar melhor compreensão do estado atual das pesquisas sobre qualidade percebida em serviços no Brasil, de grande utilidade para pesquisadores e gestores da área, o trabalho possa servir também como ponto de debate e orientação futuros para estudos empíricos que optem pela utilização de escalas existentes para mensuração de construtos de interesse.

O artigo está organizado da seguinte forma: após essa introdução, apresenta-se brevemente a importância do uso de escalas em marketing, focando em seguida, dentro da literatura de gestão de serviços, o tema de qualidade percebida dos serviços e, particularmente, as escalas SERVQUAL e SERVPERF. Após o detalhamento da metodologia, apresentam-se os resultados do estudo e é feita uma discussão sobre a adequação do uso das escalas SERVQUAL e SERVPERF em pesquisas empíricas sobre qualidade percebida em serviços no Brasil. Concluindo, são apresentadas as considerações finais do estudo.

REFERENCIAL TEÓRICO

Nas ciências sociais, em geral, e na administração, em particular, construtos são definidos como fenômenos não observáveis, que não podem ser medidos diretamente, mas que podem ser inferidos a partir de variáveis observáveis. Desta forma, o trabalho de delinear o domínio de um construto busca apreender a essência de sua natureza. Porém, os pesquisadores devem atentar ao fato de que, no caso de construtos mais “comuns” ou “populares”, podem existir múltiplas definições e conceitos na literatura, eventualmente com escopos distintos e, até mesmo, conflitantes (Spector, 1992). Por isso, é importante que o pesquisador, ao estudar um construto, faça uma avaliação cuidadosa “do que está incluído na definição e do que está excluído” (Churchill, 1979, p.67).

A medição de construtos, em pesquisas científicas na área de administração, inclusive em marketing, é feita comumente por meio de escalas. Há diversas escalas já validadas que são amplamente utilizadas (Bearden & Netemayer, 1999). No Brasil, Vieira (2011) fez uma compilação de escalas de marketing utilizadas em pesquisas no país. Há vantagens e riscos na adoção de escalas já existentes, testadas e validadas. Os pesquisadores podem comparar resultados obtidos com outras pesquisas que adotaram a mesma escala, apesar de haver o risco de se replicar escalas a contextos distintos. Adotar uma escala já testada e avaliada exige menos esforços e traz menos riscos ao pesquisador do que desenvolver uma escala nova, especialmente nos casos em que o objetivo principal do pesquisador é examinar relações entre construtos e não construir escalas. Os riscos podem envolver a não aderência plena da escala ao construto de interesse do pesquisador, ou ao contexto do estudo, a dificuldade de traduzir escalas originalmente elaboradas em outros idiomas, devido a diferenças de idiomas e culturas, e a necessidade de realizar modificações e adaptações realizadas na escala original. A não observância e inapropriada gestão de tais riscos podem levar a resultados de pesquisa que não retratam o construto de interesse do pesquisador, impactando negativamente seus achados.

Qualidade Percebida dos Serviços

O debate sobre a questão da qualidade ganhou relevância na área de administração a partir da segunda metade do século XX. Apesar do foco inicial estar na gestão da qualidade total para ambientes industriais, autores seminais, como Juran (1998), já apontavam aspectos intangíveis e subjetivos, realçando a necessidade de melhor compreensão das expectativas e percepções nas avaliações de qualidade feitas pelos clientes.

O crescimento do setor de serviços estimulou, sobretudo a partir dos anos 1980, novos estudos para identificar aspectos mais subjetivos da qualidade e seus impactos nas organizações

prestadoras de serviços, corroborando a necessidade de entendimento das percepções dos consumidores sobre a qualidade no consumo (Holbrook & Corfman, 1985). Tais estudos focam o julgamento de caráter mais abstrato, feito pelos clientes, acerca de eventual superioridade ou excelência de um produto ou serviço consumido (Zeithmal, 1988; Castleberry & McIntyre, 2011).

As percepções de qualidade são ainda mais relevantes para serviços, dada sua natureza subjetiva, associada a atributos, como intangibilidade, heterogeneidade e inseparabilidade (Grönroos, 1995; Parasuraman, Zeithmal, & Berry, 1985). A intangibilidade faz parte de sua essência, uma vez que serviços são, de forma geral, intangíveis, embora possam ter características tangíveis. Já a inseparabilidade significa que boa parte dos serviços precisa ser consumida no momento de sua produção. A variabilidade se refere a que há mais dificuldade em oferecer serviços homogêneos, como no caso de produtos, pois aquele que presta o serviço pode variar, e, assim, também a forma pela qual o faz. Além disso, tem-se a perecibilidade, que implica que a maior parte dos serviços não pode ser estocada. Por tudo isso, a medição da percepção que o cliente tem da qualidade do serviço assume particular importância.

A avaliação da qualidade percebida é fundamental para a gestão das empresas de serviços e está associada a seu potencial de impacto em diversas outras áreas de comportamento do consumidor, como satisfação, intenções de compra, lealdade e recomendação (Cronin & Taylor, 1992; Brady *et al.*, 2002; Dalmonech *et al.*, 2016). Percepções de superioridade relativa de qualidade permitem às empresas gerar vantagens competitivas, cobrar preços *premium* e ganhar participação no mercado, tendo sido sua relação positiva com a rentabilidade encontrada empiricamente (Zeithmal *et al.*, 1990). Sendo assim, a melhoria da medição das percepções de qualidade dos clientes deve ser parte da agenda estratégica e operacional das organizações.

As escalas SERVQUAL / SERVPERF

Apesar de sua relevância, a qualidade percebida em serviços, por sua própria natureza, apresenta significantes desafios para sua mensuração, o que levou diversos autores a se interessarem pelo tema (Bolton & Drew, 1991; Boulding *et al.*, 1993; Cronin & Taylor, 1992, 1994; Parasuraman *et al.*, 1985, 1988, 1991, 1994). Recentemente, outros autores têm buscado desenvolver escalas para aplicação em situações e contextos peculiares (como o trabalho de Kanyurhi, 2017, que desenvolve uma escala para aplicação em organizações de microfinanças) ou modificando escalas existentes para uso em situações específicas (como é o caso do estudo de Ali & Raza, 2017, em bancos islâmicos), para citar apenas alguns.

Com relação ao construto qualidade percebida, ao observar a produção acadêmica na área desde 1950, Abreu *et al.* (2019) apontam a relevância das obras de Parasuraman, Zeithmal e Berry, responsáveis pelo desenvolvimento da escala SERVQUAL, instrumento dominante desde a década de 1980. Em sua pesquisa bibliométrica, Abreu *et al.* (2019) observam que os artigos de Parasuraman, Zeithmal e Berry aparecem nas três primeiras posições de citações entre os 11.890 documentos que utilizam os termos *quality* e *service* na base World of Science (WOS). Destaca-se, também, na quarta posição, a obra de Cronin e Taylor (1992), autores responsáveis pelo desenvolvimento da escala SERVPERF. Mesmo havendo escalas mais recentes, as duas seguem entre as mais utilizadas. Sua ampla aceitação e volume de utilização tornam-se os dois principais fatores para sua eleição por este trabalho, visto que tal escolha busca não somente identificar escalas previamente validadas e existentes, mas também com ampla produção nacional de artigos empíricos baseados em sua aplicação prática.

A escala SERVQUAL se apoia na Teoria de Desconfirmação de Expectativas (Oliver, 1997) e sugere que a qualidade percebida de um serviço seja resultado dos hiatos (*gaps*) entre as expectativas e a performance percebida pelos clientes (Parasuraman *et al.*, 1985). Essas percepções de qualidade podem ser mensuradas por meio de 22 variáveis para avaliação das expectativas e outras 22 similares para a performance, divididas em cinco dimensões distintas e

comuns aos diversos tipos de serviços: Tangibilidade, Confiabilidade, Responsividade, Garantia / Segurança e Empatia (Parasaruman *et al.*, 1988, 1991). Para medição das variáveis, utiliza-se escala Likert de 7 pontos, variando entre “concordo plenamente” (7) a “discordo plenamente” (1).

Cronin e Taylor (1992, 1994) conceituam a qualidade percebida como uma atitude e entendem que ela pode ser capturada de forma mais adequada com base em modelo de adequação- importância. Dessa forma, entendem que a qualidade percebida do serviço deva ser mensurada por meio da avaliação da performance percebida pelo consumidor, não necessitando averiguar suas expectativas e tampouco o hiato entre essas e a performance. Os autores testam sua proposta e concluem que, dentre as 22 variáveis relacionadas à performance percebida da SERVQUAL, 21 são válidas para, em uma única dimensão, servirem como instrumento de mensuração da qualidade percebida dos serviços (Cronin & Taylor, 1992). Posteriormente, assumem que a questão da dimensionalidade pode variar (Cronin & Taylor, 1994), dependendo da natureza do serviço avaliado e sugerem, assim como Parasuraman *et al.* (1988, 1991), que a escala possa ser adaptada, dependendo das especificidades do objeto de estudo.

METODOLOGIA

Nesta seção, apresenta-se a metodologia adotada no presente estudo. Primeiramente, oferece-se uma visão geral dos passos seguidos para avaliar as escalas para mensuração do construto “qualidade percebida em serviços” nos trabalhos selecionados, a partir de um protocolo previamente elaborado. Em seguida, indicam-se os principais elementos do processo de avaliação utilizado.

Protocolo de Avaliação

Os passos para avaliação seguidos no presente trabalho foram estruturados com base na literatura sobre desenvolvimento e validação de escalas, principalmente Churchill (1979), DeVellis (2003), Netemeyer *et al.* (2003) e Spector (1992), e na experiência de dois autores do presente trabalho. A Tabela 1 apresenta o protocolo utilizado, com as etapas seguidas e seu detalhamento.

Tabela 1 – Protocolo de Avaliação para Utilização de Escalas Existentes

Etapas	Conteúdo das Etapas
Etapa 1	Especificação do Domínio do Construto <ul style="list-style-type: none"> ● Desenvolvimento da base teórica e conceitual do construto ● Revisão bibliográfica ● Avaliação do construto de interesse na pesquisa <i>vis-à-vis</i> o construto original da escala escolhida
Etapa 2	Eleição das Variáveis <ul style="list-style-type: none"> ● Identificação e escolha de escala existente na literatura ● Procedimentos de tradução da escala ● Modificações na escala: edição das variáveis originais ● Modificações na escala: adição de variáveis ● Modificações na escala: exclusão de variáveis
Etapa 3	Coleta de Dados e Pré-Teste <ul style="list-style-type: none"> ● Realização de pré-teste da escala modificada ● Avaliação das mudanças realizadas (amostra de respondentes semelhante à população-alvo)
Etapa 4	Validação da Escala <ul style="list-style-type: none"> ● Realização de análises estatísticas para dimensionalidade, confiabilidade e validade.

Etapa 5	Coleta de Dados
	<ul style="list-style-type: none"> ● Realização da coleta final de dados ● Avaliação do formato de medição utilizado
Etapa 6	Avaliação de Confiabilidade
	<ul style="list-style-type: none"> ● Realização de testes de confiabilidade e dimensionalidade com a coleta final
Etapa 7	Avaliação de Validade
	<ul style="list-style-type: none"> ● Realização de testes de validade com a coleta final
Avaliação Final	<ul style="list-style-type: none"> ● Avaliação geral, consolidando os achados das sete etapas ● Definição do nível final de adequação do uso das escalas existentes nos artigos avaliados (satisfatório; suficiente; insatisfatório)

As etapas sugeridas seguem a proposta de Churchill (1979) para a criação de novas escalas. Como os artigos deveriam utilizar escalas já previamente validadas, entendeu-se que, quando as etapas 1 e 2 forem adequadamente performadas, não há necessidade de avaliação das Etapas 3 a 7. Eventuais utilizações conceituais inadequadas do construto (Etapa 1) e/ou modificações significativas, como inclusão de novas variáveis em volume superior a 10% das variáveis originais da escala e/ou exclusão de mais que 10% das variáveis originais (Etapa 2), apontam para o risco da escala utilizada no artigo em processo de avaliação não refletir adequadamente a escala original (SERVQUAL ou SERVPERF). Nestes casos, torna-se necessária uma nova validação, por parte do pesquisador, do seu instrumento de pesquisa, sendo, por conseguinte, fundamental proceder com as avaliações das Etapas 3 a 7.

Processo de Avaliação

Detalha-se a seguir o processo de avaliação seguido no presente trabalho, incluindo a seleção dos trabalhos avaliados, o papel dos avaliadores, e como foi realizada a avaliação.

Seleção dos trabalhos avaliados

Os trabalhos foram selecionados de acordo com o escopo estabelecido no objetivo do presente trabalho, qual seja, avaliar a adequação do uso da SERVQUAL e da SERVPERF em artigos científicos brasileiros de natureza descritiva e quantitativa.

O primeiro passo foi a escolha de uma ferramenta de pesquisa dos trabalhos. Os autores selecionaram a base SPELL da ANPAD, por ser uma biblioteca eletrônica focada na “indexação, pesquisa e disponibilização gratuita de produção científica, particularmente das áreas de Administração Pública e de Empresas, Contabilidade e Turismo” (Spell, 2020). Nesta base, foram localizados os trabalhos sobre a temática escolhida, publicados nos últimos dez anos, utilizando as palavras-chaves “SERVQUAL” e “SERVPERF” e foram coletados também a classificação Qualis (CAPES) de cada um deles. A Tabela 2 apresenta os artigos obtidos nessa busca segundo a classificação dos periódicos em que foram publicados. Observe-se que não foram encontrados artigos em periódicos A1 (já que não existem periódicos brasileiros com esta classificação na área de administração), ou em periódicos classificados como A2.

Tabela 2 – Distribuição dos artigos segundo classificação Qualis (CAPES) dos periódicos

Classificação Qualis	B1	B2	B3	B4
Nº de Artigos	7	13	14	2

Decidiu-se, então, avaliar os trabalhos publicados em periódicos acadêmicos com classificação a partir de B2, por se encontrarem em veículos de maior impacto. Foram pré-selecionados 20 trabalhos, que foram nomeados de A a T, a fim de preservar o anonimato dos

autores. Os 20 trabalhos pré-selecionados foram brevemente analisados para verificar se eram de natureza descritiva e quantitativa. Desses, os autores excluíram três artigos (artigos D, O e T), por não estarem dentro do escopo (descritivo e quantitativo). Ao final, foram selecionados para avaliação 17 artigos, publicados em periódicos classificados como B1 e B2.

Papel dos avaliadores

Dos quatro autores, três foram avaliadores das escalas utilizadas nos trabalhos selecionados, e um autor foi mediador. Dos três avaliadores, um ficou responsável por fazer a avaliação da Etapa 1 da avaliação (“Especificação do Domínio do Construto”), por ser o avaliador com maior conhecimento do construto “qualidade percebida em serviços”. Todos os três avaliadores avaliaram as escalas utilizadas nos dezessete artigos selecionados, seguindo os passos estabelecidos previamente no instrumento desenvolvido de acordo com o protocolo de avaliação.

As avaliações foram feitas de forma independente por cada um dos três avaliadores, com base em um instrumento elaborado em linha com o protocolo de avaliação da pesquisa, contendo: sete perguntas para a Etapa 1; 16 perguntas para a Etapa 2; quatro para a Etapa 3; seis para a Etapa 4; quatro para a Etapa 5; seis para a Etapa 6; três para a Etapa 7; e três últimas perguntas para a Avaliação Final.

Uma vez realizadas, as notas para cada etapa, incluindo a “Avaliação Final”, foram comparadas entre os três avaliadores. Em caso de diferenças nas notas, os avaliadores discutiram até chegar a um consenso. E, em caso de dúvida a respeito de algum ponto ou falta de consenso, o mediador realizaria o desempate, discutindo com os três avaliadores. No entanto, a mediação foi desnecessária, uma vez que os avaliadores chegaram a um consenso em todos os casos examinados, demonstrando solidez do instrumento elaborado, visto que diferentes avaliadores alcançaram resultados semelhantes com a utilização de tal instrumento.

Características da avaliação

Em primeiro lugar, os avaliadores levaram em consideração se a escala utilizada foi disponibilizada nos trabalhos avaliados. Nos casos em que as escalas utilizadas foram inteiramente disponibilizadas no trabalho (seja no seu corpo, como tabela, ou como apêndice), os avaliadores puderam seguir as etapas previstas no protocolo de avaliação na íntegra. Por outro lado, nos casos em que as escalas não foram disponibilizadas (nem sequer partes dela), somente foi possível avaliar a Etapa 1 (Especificação do Domínio do Construto) destes trabalhos, não sendo possível chegar a uma avaliação final.

Para os trabalhos avaliados que somente incluíram as escalas utilizadas parcialmente, os avaliadores analisaram se havia informação suficiente para realizar a avaliação completa. Nos casos em que a escala foi incluída parcialmente, porém de forma considerada suficiente, tanto em termos de volume quanto da qualidade e consistência dos dados disponibilizados, os avaliadores realizaram todos os passos da avaliação, porém realçando a ausência de parte das variáveis do questionário com uma “ressalva” na Avaliação Final do artigo. De forma similar ao procedimento adotado quando não foi apresentada a escala utilizada, nos casos em que os avaliadores consideraram que os dados parciais das escalas inseridas nos trabalhos avaliados eram insuficientes, somente a Etapa 1 (Especificação do Domínio do Construto) da avaliação foi realizada e não houve uma avaliação final para esses artigos.

Para todos os artigos, ao final de cada etapa avaliativa, os avaliadores puderam dar notas de 1 a 5, que foram divididas em três grupos: trabalhos com notas 1 ou 2 foram classificados como “insatisfatórios”; aqueles que receberam nota 3 foram considerados “suficientes”; por último, artigos que receberam notas 4 ou 5 foram classificados como “satisfatórios”. Após as sete etapas de avaliação, os avaliadores, com base nas notas anteriores, fizeram uma avaliação geral (Avaliação Final) da utilização da escala SERVQUAL ou

SERVPERF para cada um dos artigos, seguindo o mesmo critério de medição de 1 a 5 pontos mencionado anteriormente.

Como resultado das avaliações finais, os trabalhos “insatisfatórios” são aqueles para os quais, segundo o protocolo de avaliação utilizado, não pode ser considerada “adequada” a aplicação das escalas analisadas e do construto “qualidade percebida em serviços”. Os trabalhos “suficientes”, por sua vez, são aqueles que apresentam uma consistência mínima (segundo o protocolo de avaliação utilizado) de adequação no uso das escalas e de medição do construto. Por fim, os trabalhos “satisfatórios” são aqueles que fazem uso das escalas analisadas e retratam o construto de forma adequada.

RESULTADOS

Os 17 artigos científicos publicados em periódicos de classificação Qualis B1 e B2, objeto de avaliação na pesquisa, são nomeados de A a T. Nesses artigos, observa-se a preponderância da escala SERVQUAL, que foi utilizada, individualmente, em 13 artigos e, de forma conjunta com a SERVPERF, em um artigo. Já a escala SERVPERF foi escolhida como instrumento único de mensuração da qualidade percebida em três artigos. Percebe-se que tanto o tema, quanto as escalas seguem mantendo sua relevância para as pesquisas em marketing, com a distribuição das publicações relativamente estável ao longo do período avaliado. Os dados da Tabela 3, adicionalmente, sugerem intensificação nas publicações a partir de 2015.

Tabela 3 – Distribuição anual dos artigos publicados na amostra selecionada

Ano	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Nº de Artigos	1	2	1	0	3	2	1	3	2	1

A amostra de artigos sugere, também, a capacidade de utilização e aderência destas escalas aos distintos ramos de serviços (Ladhari, 2009; Silva & Urdan, 2016). Os 17 artigos encontram-se distribuídos em dez setores diferentes, com base na Classificação Central de Produtos Básicos das Nações Unidas (Ministério da Economia, 2020), com destaque para o setor de serviços de turismo e relacionados, citado em cinco artigos. Os outros nove setores contemplam apenas um ou dois artigos entre os analisados.

De acordo com os procedimentos metodológicos adotados, a efetiva avaliação da utilização de uma escala já testada necessita de acesso, por parte dos avaliadores, aos detalhes do questionário utilizado em cada artigo. Desta forma, como etapa preparatória e fundamental para a avaliação, foi necessário checar a disponibilização do questionário utilizado por cada pesquisa. No caso dos artigos da amostra, os artigos C, E e L não forneceram qualquer detalhe sobre os itens utilizados em seus questionários e, por essa razão, não puderam ser plenamente avaliados. Dois outros artigos, F e R, apresentaram algum nível de informação sobre seus questionários no corpo do artigo, mas os avaliadores consideraram que o volume de informação não foi suficiente para que as análises pudessem seguir adiante.

Dentre os doze artigos restantes, nove publicaram seus questionários integralmente, seja no corpo do artigo ou em forma de apêndice. Os três artigos restantes, B, J e N, apesar de não apresentarem seus questionários na íntegra, forneceram material em quantidade e qualidade de informação consideradas pelos avaliadores como suficiente para que uma avaliação geral, mesmo com ressalvas, pudesse ser feita.

Indiferente à disponibilização do questionário ou não, todos os 17 artigos foram avaliados na etapa 1 (“Especificação do Domínio do Construto”) e, parcialmente, na etapa 2 (“Eleição das Variáveis”), visto que tais etapas tratam, sobretudo, de processos conceituais associados à utilização de uma escala e, por serem anteriores à própria escolha de uma escala, podem ser avaliados em linha com o protocolo de avaliação pré-estabelecido nesse trabalho.

Etapa 1 – Especificação do Domínio do Construto

De forma geral, todos os 17 artigos procederam de forma adequada na eleição e comunicação do construto associado às duas escalas da pesquisa, sendo que quinze foram conceituados como “satisfatórios” e somente dois com conceitos “suficiente”. Em todos os artigos, o objetivo de pesquisa estava associado ao construto de qualidade percebida em serviços, aspecto fundamental para a adequada utilização das escalas SERVQUAL e SERVPERF.

Atributos relevantes e desejáveis relacionados às propostas dos autores de cada escala para o construto foram, no mínimo, parcialmente considerados por todos os pesquisadores. Destaca-se a observação do conceito de hiato ou *gap* entre expectativas e performance como fator gerador para a qualidade percebida, item mencionado por todos os autores que utilizaram a escala SERVQUAL. As características de subjetividade e as dificuldades de mensuração também foram citadas pela maioria dos trabalhos, usassem SERVQUAL ou SERVPERF.

Com relação à revisão bibliográfica associada à escala, nota-se que, para quase a totalidade do grupo de artigos baseados na SERVQUAL, os autores mencionam a obra seminal de Parasuraman *et al.* (1988) como sua fonte de pesquisa. Apesar deste artigo haver sido a base para a divulgação da SERVQUAL, os mesmos autores escreveram, porém, em 1991, uma revisão do texto de 1988. Apesar de não apresentarem alterações na natureza de suas 22 variáveis, propuseram alterações de edição em diversas delas, já que a formatação do questionário original poderia gerar níveis de expectativas irrealistas e os respondentes poderiam ser confundidos pelo formato negativo de algumas frases (Parasaruman *et al.*, 1991). No entanto, apenas dois artigos mencionaram esta obra revista como fonte de seus trabalhos, assim como apenas o artigo L menciona o livro escrito posteriormente pelos mesmos autores (Zeithmal *et al.*, 1990), que também incluí revisões da escala de 1988.

Com relação à bibliografia adicional, os autores fazem uso de, em média 3,4 referências vinculadas aos autores das duas escalas, promovendo uma cobertura, no mínimo parcial, de fontes bibliográficas relevantes no processo de desenvolvimento do construto e das escalas. Esta cobertura bibliográfica é avaliada como suficiente ou satisfatória, de acordo com as premissas utilizadas pelos avaliadores. Além dos textos seminais da SERVQUAL (Parasuraman *et al.*, 1988) e da SERVPERF (Cronin & Taylor, 1992) entre as referências utilizadas, destaca-se a obra de Parasuraman *et al.* (1985), mencionada por 14 artigos, que trata dos fundamentos da qualidade percebida em serviços, enquanto fruto de hiatos (*gaps*) e serve como base para seus trabalhos futuros sobre a SERVQUAL. Outras obras de Zeithmal, principalmente seu artigo de 1988 sobre as características inerentes da percepção de qualidade em serviços (Zeithmal, 1988), apareceram em sete artigos, e de Grönroos (1995), citado em oito artigos, também merecem destaque.

Por outro lado, apesar de as escalas terem como características a possibilidade de aceitar adaptações para as áreas de serviços nas quais são aplicadas, não foram encontradas evidências de análises bibliográficas de eventuais especificidades do construto para o setor específico analisado em seis dos 17 artigos, indicando uma eventual lacuna na revisão da literatura pertinente.

Etapa 2 – Eleição das Variáveis

A etapa seguinte consiste na avaliação da adequação do processo de eleição de variáveis por parte dos autores dos artigos analisados. Como as escalas originais já haviam sido devidamente validadas, o foco da análise recai sobre o adequado respeito à estrutura original da escala e, em caso de alterações, do esclarecimento do porquê de tais modificações, apoiando-se na teoria e nos estudos empíricos existentes na literatura. A Tabela 4 apresenta os resultados da Etapa 2 da avaliação.

Tabela 4 – Adequação dos Artigos à Etapa 2 – Eleição das Variáveis

Adequação	Não se Aplica	Insatisfatória	Suficiente	Satisfatória
Nº de Artigos	5 (C/ E/ F/ L/ R)	4 (A/ B/ M/ Q)	4 (G/ H/ K/ P)	4 (I/ J/ N/ S)

Observa-se que a ausência de informações suficientes sobre os questionários utilizados e, por conseguinte, sobre as variáveis de pesquisa, impossibilita a avaliação geral de cinco dos artigos da amostra (C, E, F, L e R). Assim mesmo, parte das subavaliações da etapa ainda podem ser realizadas e estão refletidas nas análises posteriores mencionadas nessa etapa.

Entre os doze artigos restantes, quatro foram avaliados como cumprindo uma insatisfatória eleição das variáveis. Em todos os quatro artigos assim avaliados, não foram encontradas quaisquer evidências do processo de tradução utilizado pelos autores, seja próprio ou feito por terceiros, tanto para a escala SERVPERF quanto para a SERVQUAL. Adicionalmente, três deles incluem variáveis adicionais sem justificativas ou com justificativas sem embasamento teórico associado ao construto. Há exclusão de variáveis em todos os quatro, sendo que apenas um apresenta alguma argumentação relacionada à adaptação do construto à área do estudo. Desta forma, as variáveis eleitas pelos autores não mantêm ou, no caso de dois artigos, mantêm apenas parcialmente a compreensão da escala original. Por fim, constatou-se que, em todos os quatro artigos, a inclusão de novas variáveis e/ou a exclusão de variáveis originais ultrapassou o limite de 10%, utilizado neste trabalho. Por conseguinte, entende-se que os autores optaram pela formatação de uma nova escala e, por conseguinte, devem passar por todos os passos de validação da escala, de modo a garantir o alinhamento do construto qualidade percebida com seus achados.

Os artigos G, H, K e P foram classificados como “suficientes” na eleição de suas variáveis. Apesar de conseguirem uma sólida adequação nas especificações dos domínios de construtos, esses artigos tampouco atingem o limite de 10% ao incluírem e/ou excluïrem variáveis acima do nível adequado estabelecido neste trabalho para a manutenção das escalas originais. Por conseguinte, devem proceder com as etapas posteriores para a validação de suas escalas utilizadas. Entretanto, diferentemente do grupo avaliado como insatisfatório, observa-se: (i) a menção da utilização de métodos de tradução nos artigos G e K; (ii) a apresentação de justificativas, como revisão bibliográfica e consulta com especialistas, para a inclusão ou exclusão de variáveis em todos os artigos; e (iii) a manutenção da compreensão das variáveis originais mantidas. Adicionalmente, nota-se que, enquanto o artigo H optou, deliberadamente, por gerar uma ampla adaptação das escalas originais, os artigos G e P cometeram um erro conceitual ao utilizar a escala SERVQUAL, porém como sendo a SERVPERF, ou seja, excluïndo as 22 variáveis de expectativas na mensuração da qualidade percebida, sendo essa a razão para que tenham, no ponto de vista dos avaliadores, que passar pelas demais etapas de avaliação e comprovar a validade estatística de seus achados. Quanto ao artigo K, sua adição de novas variáveis ultrapassou o limite de 10%, causando um risco de alteração conceitual da escala original utilizada no artigo, tornando a validação da nova escala gerada necessária para garantir a correta aderência em seu estudo ao construto qualidade percebida.

Os artigos I, J, N e S apresentam uma sólida manutenção, tanto qualitativa como quantitativa, das bases das escalas originais, SERVQUAL e SERVPERF. Adicionado aos seus resultados consistentes da etapa 1 (Especificação do Domínio do Construto), entende-se que esses quatro artigos cumpram com as exigências de adequação para a utilização de escalas previamente validadas e, por conseguinte, nenhuma análise adicional é necessária por parte dos avaliadores. Deve-se ressaltar, porém, que os artigos J e N não divulgam integralmente a descrição das variáveis de sua escala. Assim mesmo, o volume e a consistência dos itens apresentados levam os avaliadores a validar, com ressalvas, sua adequação no uso da escala

SERVQUAL.

Por último, em uma análise dos subitens da Etapa 2, percebe-se que: (i) todos os artigos justificam adequadamente suas escolhas das escalas SERVQUAL / SERVPERF; (ii) apenas dois artigos fazem menção a processos de tradução da escala, lembrando-se que ambas as escalas são, originalmente, formatadas em língua inglesa; (iii) entre os doze artigos com disponibilização das variáveis, oito optam por incluir variáveis adicionais, porém dois não justificam suas opções; e (iv) nove também excluem itens da escala original, porém não foram encontradas justificativas em três casos.

Etapas 3 e 4 – Pré-teste e Validação da Escala

Entre os oito artigos que ultrapassam o limite de 10% e/ou que não mantêm uma adequada compreensão da escala original, apenas o artigo H realiza a etapa de Pré-teste. A etapa de validação da escala, porém, não é cumprida por nenhum dos artigos.

Etapa 5 – Coleta de Dados

A etapa 5, de coleta de dados, mesmo não necessariamente sendo utilizada pelos pesquisadores para as etapas finais de testes de confiabilidade e validade das “novas” escalas dos oito artigos, é realizada por todos eles devido ao caráter descritivo das pesquisas elencadas. Assim mesmo, a aderência aos procedimentos para uma adequada coleta de acordo com as escalas escolhidas resulta em: quatro artigos com avaliações insatisfatórias (A, M, Q e P); três artigos com avaliações suficientes (B, G e K); e apenas um artigo com avaliação satisfatória (H). A não aplicação do modelo de hiatos (*gaps*) foi a principal razão para a inadequação dos quatro artigos ao atendimento do formato de aplicação das escalas originais. Também se observou, nos artigos M, P e Q, a alteração da medida de mensuração proposta na escala original, não havendo claras justificativas para tal em mais de 50% desses casos.

Etapa 6 – Avaliação de Confiabilidade

Em três artigos (A, B e Q), não é encontrado nenhum comentário, ou apenas comentários incompletos, sobre testes de confiabilidade e dimensionalidade das escalas, gerando avaliações insatisfatórias. A avaliação é satisfatória nos demais cinco artigos, com a observação de comentários sobre os testes de confiabilidade e dimensionalidade executados, incluindo a especificação de quais foram utilizados, bem como a menção de sua aplicação a nível de dimensão.

Etapa 7 – Avaliação de Validade

Comentários sobre a aplicação e especificação dos testes de validade não são encontrados em quatro (A, B, M e Q) dos oito artigos. Nos restantes, três (G, H e K) foram avaliados como satisfatórios e um como suficiente (P).

Etapa 8 – Resultado Geral

Com base nos resultados obtidos nas etapas preliminares, com base no protocolo de avaliação e demais procedimentos adotados no presente trabalho, observou-se que: (i) cinco artigos não puderam ser completamente avaliados por insuficiência de dados sobre as variáveis de suas escalas (C, E, F, L e R); (ii) quatro artigos foram avaliados como “insatisfatórios”, não logrando manter a proposta original das escalas SERVQUAL e SERVPERV utilizadas (A, B, M e Q); (iii) quatro artigos foram avaliados como “suficientes” com as ressalvas de que, em dois casos (J e N), não foi possível obter o total detalhamento de todas as variáveis utilizadas e, nos outros dois (G e P), o não atendimento total dos métodos de aplicação da escala original, SERVQUAL, foi atenuado pelos testes estatísticos de confiabilidade e validade executados no decorrer dos trabalhos; (iv) quatro artigos (H, I, K e S) tiveram avaliações “satisfatórias”,

indicando uma adequada utilização das escalas escolhidas. A tabela 5 resume os resultados da aplicação do protocolo de avaliação aos 17 artigos selecionados neste trabalho.

Tabela 5 – Síntese dos Resultados por Artigo e Etapa de Avaliação

Artigos	C/ E/ F/ L/ R	A/ B/ M/ Q	G/ P	J/ N	H/ I/ K/ S
Disponibilização das Variáveis	Indisponível	Integral ou suficiente	Integral	Suficiente	Integral
Etapa 1	SU (C) SA (E, F, L, R)	SA (todos)	SU (P) SA (G)	SA (todos)	SA (todos)
Etapa 2		I (todos)	SU (todos)	SA (todos)	SA (I, K, S)/ SU (H)
Etapa 3		I (todos)	I (todos)		I (K)/ SA (H)
Etapa 4		I (todos)	I (todos)		I (H, K)
Etapa 5		I (A, M, Q)/ SU (B)	I (P)/ SU (G)		SU (K)/ SA (H)
Etapa 6		I (A, B, Q) / SA (M)	SA (todos)		SA (H, K)
Etapa 7		I (todos)	SU (P)/ SA (G)		SA (H, K)
Avaliação Geral	Não avaliado	I (todos)	SU (todos)	SU (todos)	SA (todos)
Ressalva			sim	sim	

Notação: I – Insatisfatória; SU – Suficiente; SA - Satisfatória

Os artigos C, E, F, L e R não disponibilizam um conjunto mínimo de informações acerca de suas variáveis e, portanto, não podem ser avaliados com base no protocolo de pesquisa adotado.

Os artigos, A, B, M, Q, apresentam, no todo, resultados que tornam suas aplicações da SERVQUAL e SERVPERF insatisfatórias. Observa-se que, mesmo alcançando níveis satisfatórios na Etapa 1, cada artigo performa de maneira insatisfatória na maioria das seis etapas seguintes. A extensiva alteração do formato e da compreensão das escalas originais durante a Etapa 2 e inadequada performance na maioria das etapas seguintes levam a níveis considerados insatisfatórios de utilização das escalas nos quatro artigos.

Os artigos G e P, por outro lado, mesmo apresentando níveis suficientes de manutenção das escalas originais na Etapa 2, têm suas outras etapas avaliadas por não cumprirem o limite de 10% na inclusão de novos itens e/ou exclusão de variáveis originais. Porém, diferentemente dos artigos classificados como insatisfatórios, os dois artigos apresentam, além de uma melhor avaliação na etapa 2, resultados satisfatórios e/ou suficientes nas etapas de confiabilidade e validação das escalas. Por conseguinte, seus resultados gerais foram avaliados como suficientes, mas com a ressalva associada às alterações observadas na etapa 2.

Os artigos J e N levaram a ressalvas e, por conseguinte, obtiveram uma avaliação geral “suficiente”, por não terem apresentado a totalidade das variáveis utilizadas. Por outro lado,

o volume e a consistência dos itens apresentados permitem uma avaliação positiva desses trabalhos, porém com a ressalva da ausência da totalidade dos itens utilizados.

Os artigos I e S foram avaliados como satisfatórios, tanto na etapa 1 quanto 2. Ao cumprir o limite de 10% e ao observarem um adequado desenvolvimento da base teórica relacionado ao construto qualidade percebida em serviços, não foi necessário avaliar suas etapas posteriores, já que cumprem com os requisitos de uma adequada utilização das escalas SERVQUAL e SERVPERF.

O artigo H, porém, buscou o caminho de utilizar as escalas originais como base para a formatação de uma nova escala adaptada ao seu estudo, alterando mais que o limite de 10% proposto na avaliação e, por conseguinte, necessitou passar por uma avaliação completa. Já o artigo K, incluí mais de 10% de itens adicionais na escala e, por conseguinte, necessitou também ser avaliado em todas as etapas de avaliação deste trabalho. Observa-se que, no caso de ambos artigos, quase todas as etapas foram cumpridas de forma satisfatória. De forma unânime, os avaliadores julgaram adequada a utilização das escalas originais nestes trabalhos.

DISCUSSÃO DOS RESULTADOS

É interessante notar como, após a análise dos doze artigos restantes, há certo “equilíbrio” nos resultados, com número igual de artigos sendo categorizado como insuficiente, suficiente e satisfatório. Ao avaliar a adequação do uso de escalas existentes em artigos científicos brasileiros de natureza descritiva e quantitativa, utilizando um protocolo de avaliação pré-definido e juízes independentes, os resultados obtidos no estudo realçam o risco de erros nos resultados de uma parcela significativa das pesquisas baseadas nas escalas SERVQUAL e SERVPERF no país.

Além de um terço dos artigos avaliados serem considerados insatisfatórios quanto ao uso das escalas por eles escolhidas, percebem-se riscos adicionais na avaliação mais minuciosa das etapas de avaliação deste trabalho. Nota-se, por exemplo, que oito artigos realizaram alterações consideradas excessivas nas escalas SERVQUAL e SERVPERF e, por conseguinte, deveriam aplicar testes de confiabilidade e validade nas escalas alteradas. É muito importante notar que, porém, em nenhum desses oito artigos são localizadas menções sobre a eventual necessidade de aferição da validade das escalas modificadas, sugerindo que, para os autores, as modificações não alteraram significativamente as escalas originais escolhidas. A não efetivação generalizada das etapas de validação e pré-teste, com exceção do artigo H que realiza um pré-teste apenas para adequar a edição das perguntas aos respondentes visados, também sugere que, no entendimento dos autores desses artigos, não há violação das escalas originais.

Mesmo o cumprimento das etapas posteriores de coleta, confiabilidade e validade executada por alguns autores está apenas associado aos processos metodológicos de suas pesquisas e a eventuais testes para aferição do alinhamento das escalas originais ao ambiente do estudo. Não há, novamente, qualquer menção a testes para validação das escalas modificadas (que se comportam como novas escalas). Assume-se, assim, que todos os oito artigos não tenham atentado para os possíveis riscos associados às excessivas alterações realizadas nas escalas originais.

Por serem escalas que permitem certo nível de adaptabilidade para melhor adequação ao serviço e à cultura em que se pretende aplicar (Parasuraman *et al.*, 1988, 1991), este fato pode passar uma falsa impressão de que a inclusão e exclusão de itens seria livre, o que não é o caso. A não execução ou associação das etapas 3, 4, 5, 6, 7 do protocolo, com a necessidade de validação da escala modificada (ou nova escala) pode estar associada à falta de experiência de pesquisadores com os procedimentos necessários para criação e adaptação de escalas.

Além de não se observar o reconhecimento da necessidade de validação de escalas excessivamente alteradas, merecem destaque outros aspectos encontrados nos resultados desse trabalho. Um deles é a questão da tradução de escalas originalmente desenvolvidas e validadas

em outro idioma. Com exceção de duas menções parciais a processos de tradução, não existe, entre todos os artigos avaliados, qualquer destaque para tal processo na execução das pesquisas. Este passo, esquecido, ou não mencionado pelos autores, é de suma importância, tendo em vista que “não basta traduzir um questionário literalmente. O desafio adicional é adaptá-lo de uma forma culturalmente relevante e compreensível, mantendo o significado e a intenção dos itens originais” (Sperber, 2004, p. 124, tradução nossa). Para Cha, Kim e Erlen (2007), a não menção a procedimentos de tradução de escalas pode acontecer principalmente pelos limitados recursos para contratação de tradutores bilíngues qualificados, dificultando assim a aplicação do procedimento de tradução reserva (em inglês, *back translation*). Também é possível que, por serem escalas amplamente utilizadas em diversos estudos reconhecidos no Brasil, os autores podem haver assumido que estas já estariam validadas no idioma local.

Outra preocupação se refere à fragilidade bibliográfica de alguns artigos, que não mencionam, ou mencionam inadequadamente, obras seminais sobre o construto de interesse dos autores, a qualidade percebida de serviços. Apenas três artigos mencionam as obras originais da escala SERVQUAL, e quase a metade não apresenta o total detalhamento das escalas efetivamente utilizadas. Sabe-se que, em geral, as revistas científicas impõem limitações de espaço aos autores, mas a omissão de aspectos metodológicos e conceituais essenciais não se justifica por esse ou qualquer outro motivo. Além disso, o fato de as escalas SERVQUAL e SERVPERV já serem conhecidas na área de administração poderia levar os autores a julgar dispensável sua publicação. No entanto, dado que são efetuadas modificações na escala, essas modificações, ao menos deveriam ser publicadas.

CONSIDERAÇÕES FINAIS

O presente artigo utilizou um protocolo para avaliar o uso de escalas já existentes em artigos de administração. Tal protocolo foi aplicado para avaliar a adequação do uso da SERVQUAL e da SERVPERF, escalas existentes (e dominantes) para a mensuração do construto qualidade percebida em serviços, em artigos científicos brasileiros de natureza descritiva e quantitativa.

Alguns artigos analisados foram avaliados como insatisfatórios, o que indica trabalhos publicados em periódicos que apresentam falhas estruturais na aplicação de escalas. Os riscos, neste caso, são tanto de âmbito acadêmico quanto prático. No âmbito acadêmico, outros pesquisadores podem utilizar escalas classificadas como “insatisfatórias” pressupondo que foram validadas, retroalimentando a aplicação de instrumentos com falhas estruturais. Do ponto de vista tanto gerencial quanto acadêmico, gestores e pesquisadores podem tomar decisões estratégicas ou executar trabalhos de pesquisa baseados em resultados advindos de trabalhos com falhas que podem ter prejudicado a confiabilidade e validade de seus resultados. A avaliação, tal como feita na presente pesquisa, pode auxiliar na verificação da qualidade das escalas que se deseja aplicar, sendo uma oportunidade, inclusive, de implementação de melhores práticas.

O estudo apresenta limitações. Em primeiro lugar, o protocolo de avaliação foi aplicado a uma única temática, a qualidade percebida em serviços mensurada através das escalas SERVQUAL e SERVPERF, o que pode ter gerado algum viés específico. Afinal, trata-se de escalas amplamente utilizadas, o que pode afetar a preocupação dos que a utilizam quanto à divulgação de determinados procedimentos metodológicos. Em segundo lugar, o fato de algum procedimento não ser relatado em um artigo não significa que ele não tenha sido executado. É possível que, ao menos em alguns dos estudos aqui examinados, os autores tenham realizado determinados procedimentos, mas tenham negligenciado relatá-los. Em terceiro lugar, os autores do presente trabalho estabeleceram o limite de 10% para inclusão ou exclusão de variáveis, e esse limite é arbitrário. No entanto, esse limite é, provavelmente, excessivo, e não muito estrito. Por último, entende-se o caráter subjetivo de alguns pontos do processo

avaliativo, porém a utilização de um instrumento baseado no protocolo proposto diminuiu tal risco, fato verificado pela semelhança dos resultados obtidos por avaliações independentes realizadas neste trabalho.

Em relação a futuras pesquisas, os autores sugerem que sejam avaliadas outras escalas, além da SERVQUAL e da SERVPERF, com base no protocolo de avaliação apresentado no presente trabalho. Ademais, pode-se aplicar o protocolo em trabalhos de outras áreas, além de marketing, em outros setores e, até mesmo, em outros países. Maior detalhamento do protocolo pode também ser útil a pesquisadores e a avaliadores de artigos científicos.

REFERÊNCIAS

- Abreu, A.; Antonialli, L.; Andrade, D. (2019) “Explorando a base intelectual do tema qualidade em serviços: quando o discurso muda o tom,” *Brazilian Journal of Marketing*, 18(2), 137-168.
- Ali, M. & Raza, S.A. (2017) “Service quality perception and customer satisfaction in Islamic banks of Pakistan: the modified SERVQUAL model,” *Total Quality Management & Business Excellence*, 28:5-6, 559-577.
- Alves, C; Madeira, S; Macambira, J. “Considerações sobre a Dinâmica do Setor de Serviços Cearense: Uma Análise Sob a Ótica do Mercado de Trabalho” *Planejamento e Políticas Públicas*. IPEA. 38, 211-235
- Bearden, W.O; Netemayer, R.G. (1999). *Handbook of Marketing Scales: Multi-item measures for marketing and consumer behavior research*. 2nd. Edition. Thousand Oaks: Sage.
- Bolton, R.; Drew, J. (1991) “A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes,” *Journal of Marketing*, 55(Jan), 1-9.
- Boulding, W.; Kalra, A.; Staelin, R.; Zeithmal, V. (1993) “A dynamic process model of service quality: from expectations to behavioral intentions,” *Journal of Marketing Research*, 30(Feb), 7-27.
- Brady, M.; Cronin, J.; Brand, R. (2002) “Performance-only measurement of service quality: a replication and extension,” *Journal of Business Research*, 55, 17-31.
- Cardoso, V; Perobelli, F. (2013) “A intensidade de serviços na economia brasileira: uma abordagem de insumo-produto” *Annals of Encontro Nacional de Economia*, 1-20.
- Castleberry, S.; McIntyre, F. (2011) “Consumers’ Quality Evaluation Process,” *Journal of Applied Business Research*, 8(3), 74-82.
- Cha, E.-S., Kim, K.H. and Erlen, J.A. (2007) “Translation of scales in cross-cultural research: issues and techniques”. *Journal of Advanced Nursing*, 58: 386-395.
- Churchill, G. (1979) “A paradigm for better measures of marketing constructs,” *Journal of Marketing Research*, 16(Feb): 64-73.
- Cronin, J.; Taylor, S. (1992) “Measuring Service Quality: A Reexamination and Extension,” *Journal of Marketing*, 56(3), 55-68.
- Cronin, J.; Taylor, S. (1994) “SERVPERF versus SERVQUAL: Reconciling Performance-Based and Perceptions-minus-Expectations Measurement of Service Quality,” *Journal of Marketing*, 58(Jan), 125-131.
- Dalmonech, H.; Goularte, M.; Ramos, P.; Monte-Mor, D. (2016) “Fatores que influenciam uma indicação de uma instituição pública federal de ensino por seus alunos através do marketing boca a boca,” *Revista Eletrônica de Ciência Administrativa*, 15(1), 25-39.
- DeVellis, R. F. (2003). *Scale Development: theory and applications*. Thousand Oaks: Sage.

- Grönroos, C. (1995). *Marketing: gerenciamento e serviços*. Rio de Janeiro: Campus.
- Holbrook, M.; Corfman, K. (1985) "Quality and Value in the Consumption Experience: Phaldrus Rides Again," in *Perceived Quality: How Consumers View Stores and Merchandise*, J. Jacoby and J. Olson (eds.) Lexington: Lexington Books, 31-57.
- IBGE. (2017) "Pesquisa anual de serviços - PAS". (accessed May 17 2020), [available at https://biblioteca.ibge.gov.br/visualizacao/periodicos/150/pas_2017_v19_informativo.pdf]
- IPEA. (2006) *Estrutura e Dinâmica do Setor de Serviços no Brasil*. 1 ed. Brasília, IPEA.
- Juran, J. (1998) "*Juran's Quality Handbook*," 5a. Ed. New York, McGraw-Hill.
- Kanyurthi, E.B. (2017). "Customer satisfaction with the services of microfinance institutions: Scale development and validation," *Strategic Change*, 26(6), 563-574.
- Ladhari, R. (2009) "A review of twenty years of SERVQUAL research," *International Journal of Quality and Service Sciences*, 1(2), 172-198.
- Netemeyer, R.G., Bearden, W.O., Sharma, S. (2003). *Scaling Procedures: issues and applications*. Thousand Oaks: Sage.
- Ministério da Economia (2020). *Classificação dos setores de serviços*. (accessed 2020, May 20) [available at <http://www.mdic.gov.br/index.php/comercio-exterior/negociacoes-internacionais/217-negociacoes-internacionais-de-servicos/1942-ni-classificacao-dos-setores-de-servicos>]
- Oliver, R.L. (1997). *Satisfaction: a behavioral perspective on the consumer*. Boston: Irwin/McGraw-Hill.
- Parasaruman, A.; Zeithmal, V.; Berry, L. (1985) "A Conceptual Model of Service Quality and Its Implications for Future Research," *The Journal of Marketing*, 49(4), 41-50.
- Parasaruman, A.; Zeithmal, V.; Berry, L. (1988) A. "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, 64(1), 12-40.
- Parasaruman, A.; Zeithmal, V.; Berry, L. (1991) "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing*, 67(4), 420-450.
- Parasaruman, A.; Zeithmal, V.; Berry, L. (1994) "A Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research," *Journal of Marketing*, 58(Jan), 111-124.
- Silva, J.; Urdan, A. (2016) "Três Modelos Integrativos de Gestão de Serviços: Um Comparativo Sob Critérios Metateóricos," *REmark – Revista Brasileira de Marketing*, 15(5), 699-711.
- Spector, P. (1992) *Summated rating scale: an introduction*. Newbury Park: Sage University Paper Series.
- Spell: Objetivos. (accessed May 10 2020), [available at <http://www.spell.org.br/sobre/objetivos>].
- Sperber, A. D. (2004). "Translation and validation of study instruments for cross-cultural research". *Gastroenterology*, 126, 124–128.
- Vieira, V.A. (2011). *Escalas em Marketing: métricas de resposta do consumidor e de desempenho empresarial*. São Paulo: Atlas.
- Zeithmal, V. (1988) "Consumer Perception of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence," *Journal of Marketing*, 52(3), 2-22.
- Zeithmal, V.; Parasuraman, A.; Berry, L. (1990) *Delivering Quality Service: Balancing Customer Perceptions and Expectations*, New York: The Free Press.