

Escala de Avaliação do Líder Autêntico: Confiabilidade e Diferenças Entre Grupos

Maria Isabel de Campos¹, Fabián Javier Marín Rueda²

¹ <http://orcid.org/0000-0002-5949-1583> / Universidade São Francisco (USF), Brasil

² <http://orcid.org/0000-0001-5173-0802> / Universidade São Francisco (USF), Brasil; Centro Universitário de Brasília (CEUB), Brasil

Resumo

Medidas de Liderança Autêntica concentraram-se em escalas em modelo de heterorrelato, dificultando a avaliação de líderes em alguns processos da prática organizacional. A Escala de Avaliação do Líder Autêntico (EALA), desenvolvida e validada em formato de teste de julgamento situacional, pretende solucionar esse problema, mas é ainda carente de investigação de confiabilidade. Este estudo teve dois objetivos: investigar a confiabilidade da EALA, assim como verificar as diferenças entre grupos caracterizados por variáveis sociodemográficas. Participaram 41 profissionais brasileiros, dos quais 23 estavam exercendo funções de liderança. Os resultados indicaram bons índices de confiabilidade para a EALA, por meio de teste-reteste. Para o fator autoconsciência, foi encontrada uma diferença estatisticamente significativa entre os grupos caracterizados por estar ou não exercendo funções de liderança. Foram apresentadas análises sobre a qualidade do instrumento e sobre a teoria da Liderança Autêntica. Sugeriu-se a continuidade das investigações com a EALA por meio de estudos longitudinais.

Palavras-chave: liderança, autenticidade, comportamento organizacional.

Authentic Leader Rating Scale: Reliability and Differences Between Groups

Abstract

Authentic Leadership Measures has focused on hetero-reported model scales, making it difficult to assess leaders in some organizational processes. The Authentic Leader Assessment Scale (EALA in Portuguese) was developed and validated in a situational judgment test format, aiming to solve this problem; but it is still lacking investigation of reliability. This study had two objectives: to investigate the reliability of EALA, and to verify the differences between groups characterized by socio-demographic variables. The study included 41 Brazilian professionals, among them 23 held leadership roles. The results indicated good reliability indexes for EALA, through test-retest. For the self-awareness factor, a statistically significant difference was found between the groups characterized by occupying or not occupying leadership roles. The study presents analyses on the instrument quality and on the Authentic Leadership theory. It is suggested to continue investigations with EALA throughout longitudinal studies.

Keywords: leadership, authenticity, organizational behavior.

Escala de Evaluación del Líder Auténtico: Confiabilidad y Diferencias Entre Grupos

Resumen

Las medidas de Liderazgo Auténticas se enfocan en escalas en modelo hetero informe, dificultando la evaluación de líderes en algunos procesos de práctica organizacional. La Escala de Evaluación del Líder Auténtico (EALA), desarrollada y validada en un formato de prueba de juicio situacional, pretende resolver este problema, pero aún carece de investigación de la confiabilidad. Este estudio tuvo dos objetivos: investigar a confiabilidad de la EALA, así como verificar diferencias entre grupos caracterizados por variables sociodemográficas. Participaron 41 profesionales brasileños, de los cuales 23 ejercían funciones de liderazgo. Por medio del test-retest, los resultados indicaron buenos índices de confiabilidad para la EALA. Para el factor autoconciencia se encontró una diferencia estadísticamente significativa entre los grupos caracterizados por estar o no ejerciendo funciones de liderazgo. Fueran presentados análisis sobre la calidad del instrumento y sobre la teoría del Liderazgo Auténtico. Se sugiere continuar las investigaciones con la EALA, por medio de estudios longitudinales.

Palabras clave: liderazgo, autenticidad, comportamiento organizacional.

No âmbito da Psicologia Organizacional e do Trabalho (POT) e do Comportamento Organizacional (CO), Testes de Julgamento Situacional (TJSs) e Liderança Autêntica (LA) têm recebido atenção de pesquisadores e de profissionais da prática, desde a última década do século XX e com interesse crescente desde o início deste século. Restam, entretanto, lacunas em ambos os campos de pesquisa (Ambiel, Campos, Alves & Silva, 2015; Avolio, Wernsing, & Gardner, 2018; Campos & Rueda, 2018a, 2018b; Corstjens, Lievens, & Krumm, 2017; De Leng et al., 2017; Kaminski, Felfe, Schäpers, & Krumm, 2019; Martin-Raugh et al., 2018). Acompanhando esse contexto, o objetivo deste artigo é investigar a confiabilidade de um TJS desenvolvido para avaliar a LA, endereçando duas das lacunas existentes: a confiabilidade de TJSs e a possibilidade de avaliar a LA em processos inerentes à prática organizacional, que não se beneficiam da aplicação de instrumentos de heterorrelato, tais como seleção, e treinamento e desenvolvimento.

No contexto teórico identificam-se três perspectivas sob as quais se define a LA: 1) a intrapessoal, cujo foco é o líder e o que acontece com ele; 2) a interpessoal, que descreve a LA como relacional, sendo criada conjuntamente por líderes e seguidores; e 3) a desenvolvimental, que posiciona a LA como algo que pode ser nutrido em um líder, e não como uma característica fixa. Nessa perspectiva, a LA se desenvolve em pessoas ao longo da vida, podendo ser desencadeada por grandes eventos, como uma doença grave ou uma nova carreira (Northouse, 2018). Este artigo é orientado pela terceira perspectiva, visto ser ela a que viabilizou a operacionalização de medidas.

A LA não deve, necessariamente, ser considerada como já presente nas organizações. Um dos principais objetivos da linha teórica pautada pela terceira perspectiva é possibilitar o desenvolvimento de líderes que tenham potencial para adotar o estilo e que, ao adotá-lo, passem a influenciar seguidores para assumirem um comportamento similar, ao mesmo tempo em que, conjuntamente, continuem a gerar resultados positivos para as organizações (Avolio & Gardner, 2005; Avolio, Gardner, Walumbwa, Luthans, & May, 2004; Campos & Rueda, 2019; Gardner, Avolio, Luthans, May, & Walumbwa, 2005). O desenvolvimento desse estilo de liderança pode também ocorrer por meio da execução de intervenções a exemplo dos trabalhos de Baron (2016) e de Fusco (2018), que tratam de treinamentos e de *coaching* com o objetivo de desenvolver a LA em indivíduos.

O comportamento de um líder autêntico distingue-se por ser pautado pela ética e por suas ações se apresentarem em consonância estrita com seus valores pessoais internalizados, sempre positivos (Avolio & Gardner, 2005; Avolio et al., 2004; Avolio & Walumbwa, 2014; Gardner et al., 2005). Visando operacionalizar a LA, o que resultou na validação inicial do primeiro instrumento de medida da área - o *Authentic Leadership Questionnaire* (ALQ) -, Walumbwa, Avolio, Gardner, Wernsing e Peterson (2008) a posicionaram como um construto multidimensional de segunda ordem, formado por quatro componentes. Os componentes de primeira ordem são: a autoconsciência (AC - indica que o líder compreende suas próprias forças e limitações, sendo também capaz de perceber como elas afetam outros indivíduos, especialmente seus seguidores); o processamento balanceado (PB - indica que o líder solicita e analisa informações relevantes e pontos de vista diferentes dos seus antes de tomar uma decisão); a perspectiva moral internalizada (PMI - refere-se à autorregulação orientada por valores e padrões morais internalizados); e a transparência de relacionamento (TR - refere-se a uma abertura e disposição de compartilhar seus verdadeiros pensamentos e sentimentos, o seu *self* autêntico, em especial com seus seguidores).

A partir da validação inicial do ALQ, ocorreu uma adesão

global à realização de experimentos com base em sua aplicação, sendo o trabalho de Cervo et. al (2018) um exemplo de estudo realizado no Brasil. A literatura passou a reunir, também, críticas sobre a LA. As críticas se direcionaram tanto ao embasamento filosófico, abrangência e contribuição teórica, como também às propriedades psicométricas do instrumento e à natureza de primeira ou de segunda ordem do construto (Campos & Rueda, 2018b).

Ford e Harding (2011), por exemplo, defenderam que o líder autêntico parece ser alguém plenamente absorvido pela organização, o que acabaria por fazer dele um objeto carente de subjetividade. Nesse sentido, seu papel como líder seria assegurar que os seguidores se tornassem também objetos, com suas reivindicações de subjetividade postas à mercê dos valores do líder e da organização. Para as autoras, esse seria um relacionamento de dominação/supressão e não um relacionamento líder/seguidor.

Algera e Lips-Wiersma (2012) acrescentaram que a inautenticidade é inevitável, ao mesmo tempo em que a autenticidade requer a criação de um significado próprio do indivíduo, não implicando, necessariamente, em congruência de objetivos e valores. Principalmente, para esses autores, a autenticidade não é intrinsecamente ética.

Hopkins e O'Neil (2015) defenderam que a teoria da LA negligencia a avaliação de como ela se aplica às mulheres. Consideraram que as organizações são entidades de gênero, que requerem que mulheres se adequem a um ambiente dominado por homens e afirmaram que existe um dilema duplo inerente a essa questão, que obriga mulheres a fazerem uma escolha entre agir em concordância com os comportamentos normativos de gênero, ou com o comportamento esperado para o papel de liderança. Desta forma, o peso dado ao verdadeiro *self* individual, em oposição ao *self* em relação aos outros, continuaria a posicionar as mulheres como estranhas à liderança nos ambientes organizacionais.

Em relação às propriedades psicométricas do instrumento ALQ e à natureza de primeira ou de segunda ordem do construto LA, Neider e Schriesheim (2011) questionaram tanto a qualidade do processo usado pelos desenvolvedores da escala na busca de evidências de validade de conteúdo, como também a qualidade do instrumento aplicado para evidenciar sua validade discriminante. Os autores consideraram também, que para o progresso das pesquisas, seria importante a existência de um instrumento com todos os itens publicados, o que não ocorrera com o ALQ. Por esses motivos, desenvolveram e validaram o instrumento *Authentic Leadership Inventory* (ALI), com estrutura similar à do ALQ.

As críticas seguiram acompanhando a realização de pesquisas empíricas. Uma metanálise que contou com 100 artigos encontrou alta correlação entre a LA e a liderança transformacional ($r > 0,70$) e carência de validade incremental, indicando redundância entre os construtos (Banks, Gooty, Ross, Williams, & Harrington, 2018; Banks, McCauley, Gardner, & Guler, 2016). Um terceiro instrumento, o *Authentic Leadership Integrated Questionnaire* (AL-IQ), foi validado com a intenção de melhorar a precisão das medidas previamente existentes (Levesque-Côté, Fernet, Austin, & Morin, 2017). Antonakis (2017) questionou a existência de ideologia (e.g. o líder deve ser bom) e a inexistência de exogeneidade (i.e. a natureza do que é medido deve ser externa com relação aos resultados que promove) quando da definição da LA, o que teria levado à circularidade da teoria (e.g. o bom líder é autêntico e o líder autêntico faz o bem).

Assim, criticando o excesso de positividade abarcado pela teoria e de certa forma resumindo as críticas prévias, Alvesson e Einola (2019) afirmaram que a LA apresenta fundamentos filosóficos e teóricos instáveis, raciocínio tautológico, estudos empíricos fracos, instrumentos de medição sem sentido, afirmações sem

embasamento epistemológico e uma visão simplista a respeito da vida corporativa. É sob esse contexto que algumas propostas para reconceitualização do construto foram apresentadas (e. g. Johnsen, 2018; Kempster, Iszatt-White, & Brown, 2018; Sidani & Rowe, 2018).

Avolio et al. (2018), entretanto, consideraram que o uso dos quatro fatores da LA como propostos em 2008 ainda se justifica e que eles apresentam suficiente profundidade no que diz respeito ao que constitui o construto. Consideraram, ainda, que essa proposta é facilitadora da promoção de novos estudos, que viessem a ampliar entendimentos sobre a rede nomológica da LA e que novas investigações deveriam aplicar métodos estatísticos mais recentes, que viessem a fornecer maneiras alternativas de modelar a complexidade sistêmica da questão, percebendo-a como um fenômeno, e não como uma única fonte latente.

Nesse contexto, Campos e Rueda (2020) desenvolveram e apresentaram evidências iniciais de validade para um TJS que avalia o comportamento do líder, com base na teoria da LA, sob a perspectiva desenvolvimental. Trata-se da Escala de Avaliação do Líder Autêntico (EALA). Os autores adotaram o TJS como método de medida em função de os instrumentos de avaliação da LA até então validados (ALQ, ALI e AL-IQ) serem escalas tipo Likert em modelo de heterorrelato (líderes avaliados por seus liderados).

Segundo Campos e Rueda (2018a), o formato de teste em heterorrelato não tem potencial de aplicação em processos de seleção, treinamento e desenvolvimento, e outros inerentes à prática nas organizações, constituindo uma das lacunas de medidas no campo da LA. Ainda, segundo Weiss, Razinskas, Backmann e Hoegl (2018), embora diversos aspectos da liderança possam ser avaliados por seguidores, esse não seria o caso da LA. Para esses autores, seguidores não estariam aptos a assegurar que um líder age de forma consistente com seus pensamentos (valores e crenças) e sentimentos (emoções), o que é considerado atributo núcleo da LA.

Diferentemente de instrumentos de heterorrelato, TJSs são amplamente utilizados em processos de seleção em ambientes organizacionais ou de formação profissional. Normalmente, TJSs são bem aceitos pelos respondentes, que vêm neles uma proximidade maior com sua realidade laboral ou profissional. Ainda, sua validade preditiva e incremental têm alcançado bons resultados com relação a desempenho e outros critérios relacionados ao trabalho (Corstjens et al., 2017; Goss et al., 2017; Kaminski et al., 2019; Martin-Raugh et al., 2018; Tiffin et al., 2019).

Um TJS é comumente formado por uma situação hipotética com viabilidade de ocorrência em determinado âmbito profissional (*stem item*) e um conjunto de possíveis soluções para a dada situação (*response items*). As situações hipotéticas (cenários) são desenvolvidas com base em incidentes típicos, ou em algum outro conteúdo relacionado ao trabalho ou à profissão. Tipicamente, o respondente é convidado a selecionar soluções por meio de sua adequação, eficiência, ou similaridade com seu próprio comportamento, indicando para a dada situação, o que faria ou o que entende que deveria ser feito (*would do* ou *should do*), revelando tendências comportamentais ou arcabouço de conhecimentos (Ambiel et al., 2015; Catano, Brochu, & Lamerson, 2012; Corstjens et al., 2017).

TJSs vêm se popularizando em situações em que é necessário verificar (e/ou prever) atitudes que podem vir a ser adotadas pelo indivíduo em situações desafiadoras (Goss et al., 2017). Esse tipo de TJS é denominado como orientado a construto (Tiffin et al., 2019), também conhecido como TJS de domínio geral de conhecimento, em contraste com os que puramente avaliam conhecimento contextualizado (Corstjens et al., 2017). Pontuar um TJS orientado a construto é mais complicado do que pontuar um teste de conhecimento, porque não há, objetivamente, respostas

corretas ou incorretas (De Leng et al., 2017) e, normalmente, o autor do teste é quem define a forma de pontuação (Ambiel et al., 2015; Motowidlo, Dunnette, & Carter, 1990), o que pode influenciar diretamente na validade e confiabilidade desses instrumentos.

Apesar da ampla utilização em ambientes da prática e do vasto número de pesquisas realizadas, TJSs ainda enfrentam desafios relacionados a formas de se evidenciar sua validade e confiabilidade, visto que seu formato não é aderente à Teoria Clássica dos Testes, para a qual diversos métodos de análise estatística estão disponíveis (Sorrel et al., 2016). O índice alfa de Cronbach, por exemplo, não se mostra adequado para a estrutura dos TJSs que, em sua maioria, são compostos por itens heterogêneos (Goss et al., 2017; Sorrel et al., 2016; Tiffin et al., 2019).

Valores encontrados por meio de metanálises reunindo estudos que investigaram a consistência interna de TJSs têm variado entre 0,46 e 0,68. Investigações por meio de teste-reteste, método considerado mais adequado para o caso, ainda são escassas, chegando a obter, em alguns casos, índices um pouco melhores, com r próximo a 0,80 para TJSs que avaliam conhecimento e próximo a 0,70 para TJSs que avaliam comportamento (Catano et al., 2012; Corstjens et al. 2017).

Ao apresentar as evidências iniciais de validade para a EALA, Campos e Rueda (2020) nada abordaram a respeito da confiabilidade deste TJS. Para continuidade de pesquisas e almejando futuras aplicações práticas é relevante que essa verificação seja feita. Assim, o presente estudo teve como primeiro objetivo investigar a confiabilidade da EALA por meio de teste-reteste utilizando, para isto, novos dados coletados com uma parcela da população participante do estudo inicial. Um segundo objetivo integrou este estudo: investigar eventuais discriminações de nível de LA relativas a variáveis sociodemográficas, por meio da avaliação de diferenças de pontuação entre grupos formados por sexo, faixa etária, estar ou não exercendo função de liderança e tamanho da organização para a qual o respondente trabalhava.

Método

Participantes

O estudo contou com a participação de 41 indivíduos (24 homens), com idade média de 41,10 anos ($dp = 11,08$), residentes no sudeste (30), centro-oeste (4), nordeste (4), sul (1) ou fora (2) do Brasil. Dos 41 respondentes, 23 estavam exercendo funções de liderança. Com relação aos setores de atuação, 19 atuavam no setor de serviços, oito em serviços públicos, quatro no setor industrial, sete no terceiro setor e três não estavam trabalhando. O nível de escolaridade dos respondentes era ensino superior incompleto (4) ou completo (9) e pós-graduação completa ou incompleta (28). Com relação ao porte, 22 disseram estar trabalhando em organizações com até 100 funcionários e 16 em organizações com mais de 100 funcionários. Dos que estavam trabalhando, seis afirmaram não se reportar a nenhum líder ou chefe.

Instrumento

Escala de Avaliação do Líder Autêntico (Campos & Rueda, 2020). A EALA é um teste de julgamento situacional composto por 13 itens que avaliam comportamentos mais (EALA+) ou menos (EALA-) autênticos por parte de um líder, de acordo com a teoria da LA. Cada item é formado por um conjunto de situação (*stem item*) e quatro opções de solução para a situação (*response items*).

Em cada um dos itens os respondentes são orientados a selecionar duas das soluções, de acordo com as seguintes classifica-

ções e usando cada classificação uma única vez: pouco possível, ou impossível (indicando baixa ou nenhuma possibilidade de que o respondente se comportaria daquela forma, caso vivenciasse a situação); bastante possível, ou com certeza (indicando alta possibilidade ou certeza de que o respondente se comportaria daquela forma, caso vivenciasse a situação).

Um exemplo de item situacional da EALA é: “*Carlos é gerente de uma área de produção em uma indústria de grande porte. Dois chefes de equipe, diretamente subordinados a ele, entraram em conflito pessoal. Isso começou a influenciar negativamente no desempenho de tarefas das equipes, pois eles conduzem setores cujos resultados estão diretamente ligados entre si. Carlos conversou sobre o caso em uma reunião com seu superior e seus pares, na qual se verificou ser possível a transferência de um dos dois envolvidos para outra gerência, que está com uma posição aberta, porém para trabalhar em outro turno. Se estivesse no lugar de Carlos, o que você faria?*”.

Para este item situacional as opções de resposta (itens de resposta) são: a) Solicitaria ao departamento de RH que realocasse qualquer um dos chefes de equipe para a vaga disponível, comunicando-o diretamente sobre o fato. Porém, apenas após terem encontrado um substituto adequado para a mesma posição; b) Em uma reunião com ambos destacaria suas responsabilidades com relação aos resultados de suas equipes e daria um prazo para a solução da situação. Não mencionaria a vaga disponível, deixando a possibilidade como alternativa futura; c) Em uma reunião com ambos ressaltaria a necessidade de manter o profissionalismo, destacando suas responsabilidades com relação aos resultados das equipes. Mencionaria a vaga aberta no outro turno e que um deles pode pedir a transferência, caso queira; d) Ouviria separadamente cada chefe de equipe a respeito da situação e, de acordo com o seu entendimento a respeito do conflito, escolheria quem será transferido. Depois disso selecionaria um substituto para a posição e, então, informaria pessoalmente a decisão ao escolhido.

A EALA foi inicialmente validada por meio da aplicação de *Cognitive Diagnosis Model* (CDM). Seus autores encontraram que o modelo *Generalized Deterministic-input, Noisy-and-gate Model* (GDI-NA) apresentava a melhor adequação para os dados. Os índices de adequação se mostraram satisfatórios sendo $\text{Max}(X^2) = 11,56$ ($p > 0,05$); $\text{loglike} = -3739,97$ ($N_{\text{pars}} = 89$), $\text{SRMSR} = 0,038$, $\text{abs}(\text{fcor}) = 0,15$ ($p < 0,05$) para EALA+; e $\text{Max}(X^2) = 8,94$ ($p > 0,05$); $\text{loglike} = -3811,95$ ($N_{\text{pars}} = 69$), $\text{SRMSR} = 0,039$, $\text{abs}(\text{fcor}) = 0,14$ ($p < 0,05$) para EALA-.

Procedimentos de Coleta de dados e Cuidados Éticos

Após o projeto ser aprovado por um Comitê de Ética em Pesquisa sob o protocolo CAAE nº 05987019.3.0000.5514, convites para participar desta pesquisa foram enviados para 393 indivíduos, que haviam informado seu endereço eletrônico quando participaram da pesquisa para busca inicial de evidências de validade da EALA (Campos & Rueda, 2020). As mensagens eletrônicas foram enviadas manualmente entre 11 e 16 de agosto de 2019, 15 meses após a primeira coleta. Links individuais para fornecimento das respostas permaneceram em funcionamento por três meses. A coleta foi realizada por meio de uma ferramenta desenvolvida *ad hoc* sobre a plataforma Prospektor da empresa KNBS (www.knbs.net.br). Em uma tela inicial foram apresentadas as informações relativas à pesquisa e na sequência o Termo de Consentimento Livre e Esclarecido (TCLE). Os participantes que concordaram com o termo responderam inicialmente um questionário sociodemográfico e, na sequência, a EALA.

Procedimentos de Análise de Dados

Foi utilizado o *Statistical Package for the Social Sciences* (SPSS)

v.21. Inicialmente verificaram-se as estatísticas descritivas e foram calculadas as pontuações para a EALA com base no método de Motowidlo et al. (1990), utilizando-se as matrizes de pontuação validadas por Campos e Rueda (2020). Nesta forma de pontuação, se o respondente acerta as opções para a EALA+ e para a EALA-, recebe 2, se inverte a seleção de ambas, recebe -2, se acerta apenas uma das opções, recebe 1 e se inverte apenas uma das opções recebe -1. Em outros casos, o respondente recebe 0. Para evitar pontuações negativas foram adicionados 2 pontos a cada uma dessas atribuições. Desta forma, a pontuação da EALA pode variar entre 0 e 4.

Utilizando-se as pontuações dos 41 participantes para as duas coletas, isto é, para o estudo de validade inicial da EALA e para esta pesquisa, realizou-se então o teste *t* de Student para amostras pareadas, obtendo-se assim as correlações de Pearson e as diferenças de média entre as pontuações resultantes das duas coletas. Calculou-se também o tamanho do efeito por meio do *d* de Cohen (Cohen, 1988). Isto foi realizado para os 13 pares de itens da EALA, para os quatro pares de fatores (AC, PB, PMI e TR) e para o par da LA como construto de segunda ordem.

Os resultados obtidos nesta etapa foram utilizados para selecionar itens a serem eliminados da EALA. Após as eliminações repetiram-se os cálculos de pontuação da EALA e foi realizado novo teste *t* de Student para amostras pareadas e os tamanhos de efeito foram recalculados. Na sequência, foram verificadas as diferenças de pontuação para grupos de respondentes caracterizados por estar ou não ocupando função de liderança, sexo, idade e porte da organização em que trabalhavam.

Resultados

A Tabela 1 mostra os resultados do teste *t* de Student para as amostras emparelhadas dos 13 itens da EALA. Mostra, também, os tamanhos de efeito relativos às diferenças de pontuação obtidas para os itens.

Nota-se que para o par 5 não foi encontrada correlação significativa entre os resultados de teste e de reteste e que das 12 correlações significativas encontradas, uma foi negativa (par 7). Ainda, três das correlações significativas podem ser consideradas fracas (pares 4, 6 e 8), cinco moderadas (pares 3, 7, 10, 11 e 13) e quatro fortes (pares 1, 2, 9 e 12) (Dancey & Reidy, 2006).

Observa-se, também, que foram encontradas diferenças significativas entre as médias de teste e de reteste para sete dos 13 itens (2, 3, 4, 5, 7, 8 e 11) e que para dez dos pares as pontuações relativas ao reteste foram menores do que as pontuações para o teste. O tamanho do efeito se mostrou insignificante para os pares 9, 10 e 12; grande para o par 5; muito grande para o par 7; e pequeno para os demais pares (Espírito-Santo & Daniel, 2015). A Tabela 2 mostra os resultados do teste *t* de Student para as amostras emparelhadas dos fatores da EALA.

Todas as correlações encontradas foram positivas e de magnitude moderada (Dancey & Reidy, 2006). As diferenças de média encontradas entre resultados do teste-reteste só não foram significativas para transparência de relacionamento, fator para o qual o tamanho de efeito se mostrou pequeno. Para as diferenças relativas a AC, PB e fator geral LA o tamanho do efeito pode ser classificado como grande e para o fator PMI como muito grande (Cohen, 1988; Espírito-Santo & Daniel, 2015).

Os resultados apresentados na Tabela 1 permitiram concluir que os itens 5 e 7 não mostraram confiabilidade minimamente aceitável, uma vez que não houve correlação significativa para o item 5 e que no caso do item 7 a correlação foi significativa, mas negativa, indicando que houve inversão de respostas entre o teste e o reteste. O uso desses itens na computação dos resultados dos fa-

Tabela 1
Estatísticas de Amostras Emparelhadas - Teste e Reteste (Itens E.A.L.A.)

Par	T / RT	Média	dp	r	p**	Dif Média	dp	t*	p**	d
1	T	2,68	1,06	0,72	-	0,20	0,75	1,67	0,103	0,20
	RT	2,49	0,87							
2	T	2,34	0,88	0,80	-	-0,20	0,60	-2,08	0,044	0,21
	RT	2,54	1,00							
3	T	2,39	0,80	0,51	0,001	0,29	0,84	2,22	0,032	0,35
	RT	2,10	0,89							
4	T	2,54	0,90	0,36	0,022	0,34	0,88	2,48	0,018	0,45
	RT	2,20	0,60							
5	T	2,88	1,00	0,06	0,710	1,02	1,11	5,93	0,000	1,28
	RT	1,85	0,53							
6	T	3,12	1,00	0,39	0,013	0,24	1,11	1,40	0,168	0,24
	RT	2,88	1,00							
7	T	3,32	0,96	0,48	0,002	2,54	1,67	9,70	0,000	2,60
	RT	0,78	0,99							
8	T	1,95	0,31	0,35	0,024	-0,29	0,96	-1,96	0,057	0,39
	RT	2,24	1,02							
9	T	3,27	0,98	0,90	-	0,10	0,44	1,43	0,160	0,10
	RT	3,17	1,00							
10	T	2,90	0,80	0,58	-	0,12	0,87	0,90	0,376	0,13
	RT	2,78	1,04							
11	T	2,54	0,90	0,46	0,002	-0,34	0,99	-2,21	0,033	0,36
	RT	2,88	1,00							
12	T	2,44	0,92	0,88	-	0,10	0,44	1,43	0,160	0,17
	RT	2,29	0,84							
13	T	3,46	0,90	0,68	-	0,20	0,75	1,67	0,103	0,21
	RT	3,27	0,98							

Nota. * $df = 40$; ** onde ausente $p < 0,001$

Tabela 2
Estatísticas de Amostras Emparelhadas - Teste e Reteste (Fatores E.A.L.A.) - 13 Itens

Fator	T / RT	Média	dp	r	p	Dif Média	dp	t*	p	d
AC	T	3,05	0,52	0,47	0,002	0,63	0,58	6,92	-	1,12
	RT	2,43	0,60							
PB	T	2,95	0,48	0,59	-	0,58	0,43	8,78	-	1,25
	RT	2,37	0,46							
PMI	T	2,77	0,34	0,39	0,011	0,48	0,39	7,85	-	1,35
	RT	2,29	0,37							
TR	T	2,75	0,37	0,65	-	0,08	0,30	1,78	0,083	0,23
	RT	2,66	0,34							
LA	T	2,88	0,35	0,57	-	0,44	0,34	8,46	-	1,23
	RT	2,44	0,37							

Nota. * $df = 40$; ** onde ausente $p < 0,001$; AC = Autoconsciência; PB = Processamento Balanceado; PMI = Perspectiva Moral Internalizada; TR = Transparência de Relacionamento; LA = Liderança Autêntica.

tores impactou nos resultados apresentados na Tabela 2. Por esses motivos, esses itens foram eliminados da escala. A Tabela 3 mostra os resultados do teste t de Student para as amostras emparelhadas dos fatores da EALA, calculados com os 11 itens restantes.

Observa-se que com a extração dos itens 5 e 7 as correlações entre teste e reteste mostraram magnitudes maiores e que duas delas alcançaram níveis fortes (PB e fator geral LA). Ainda, as diferenças de médias encontradas deixaram de ser significativas. A exceção foi a diferença das médias de PB. O tamanho do efeito se mostrou insignificante para os fatores AC e PMI, e pequeno para TR, PB e LA geral (Cohen, 1988, Espírito-Santo & Daniel, 2015).

Tabela 3
Estatísticas de Amostras Emparelhadas - Teste e Reteste (Fatores E.A.L.A.) - 11 itens

Fator	T / RT	Média	dp	r	p	Dif Média	dp	t*	p	d
AC	T	2,97	0,54	0,61	-	-0,01	0,57	-0,09	0,927	0,01
	RT	2,98	0,70							
PB	T	2,88	0,50	0,78	-	0,20	0,32	3,86	-	0,41
	RT	2,68	0,46							
PMI	T	2,66	0,32	0,51	0,001	0,04	0,37	0,78	0,442	0,12
	RT	2,62	0,41							
TR	T	2,75	0,37	0,65	-	0,08	0,30	1,78	0,083	0,23
	RT	2,66	0,34							
LA	T	2,81	0,34	0,71	-	0,08	0,28	1,82	0,077	0,22
	RT	2,73	0,39							

Nota. * $df = 40$; ** onde ausente $p < 0,001$; AC = Autoconsciência; PB = Processamento Balanceado; PMI = Perspectiva Moral Internalizada; TR = Transparência de Relacionamento; LA = Liderança Autêntica.

Visando cumprir o segundo objetivo deste estudo, utilizando-se as pontuações com base nos 11 itens remanescentes, avaliaram-se as diferenças de média entre os grupos formados por sexo, faixa etária, estar ou não exercendo função de liderança e tamanho da organização para a qual o respondente trabalhava. A Tabela 4 mostra os resultados obtidos.

À exceção do fator autoconsciência para os grupos distinguidos pelo critério de estar ou não exercendo um cargo com funções de liderança, não foram encontradas diferenças estatisticamente significativas entre os grupos ao nível $p < 0,05$. Observa-se, entretanto, que para todos os fatores, as médias obtidas pelos respondentes que não estavam exercendo funções de liderança tiveram magnitudes maiores do que as médias dos respondentes que estavam e que para o fator transparência e para LA_Geral o valor encontrado para p foi muito próximo ao ponto de corte adotado. Para esses grupos, os tamanhos de efeito se mostraram pequenos para os fatores PB e PMI e médios para AC, TR e LA_Geral (Cohen, 1988; Espírito-Santo & Daniel, 2015).

Ressalta-se ainda que, de acordo com os critérios indicados por Cohen (1988) e por Espírito-Santo e Daniel (2015), a maioria dos tamanhos de efeitos encontrados nas análises de todos os grupos ficou na faixa classificada como insignificante, ou seja, $d < 0,20$. As exceções foram os casos dos fatores PMI e TR na análise por faixa etária e AC na análise por tamanho de empresa. Nesses três casos o tamanho de efeito se mostrou pequeno.

Discussão

Os objetivos deste artigo foram investigar a confiabilidade da Escala de Avaliação do Líder Autêntico, um teste de julgamento situacional desenvolvido para avaliar a liderança autêntica, e as diferenças entre grupos de sujeitos caracterizados por variáveis sociodemográficas. Uma análise inicial com base nas pontuações calculadas por meio dos 13 itens da EALA, de acordo com as matrizes-Q validadas por Campos e Rueda (2020) e aplicando o método de pontuação adaptado de Motowidlo et al. (1990), obteve valores de r variando entre 0,39 (fator PMI) e 0,65 (fator TR). Visando melhorar o nível de confiabilidade, dois itens (5 e 7) foram eliminados e os cálculos refeitos.

Os resultados indicaram melhoria da confiabilidade para a EALA, com $r > 0,70$ para processamento balanceado e para a LA geral, $r > 0,60$ para autoconsciência e transparência de relacionamento e $r = 0,51$ para perspectiva moral internalizada. Esses valores encontram-se nas faixas relatadas na literatura para confiabilidade de TJSs (Catano et al., 2012; Cortjens et al., 2018; Kaminski et al., 2019). Estão, entretanto, abaixo do melhor índice

Tabela 4
Diferenças de grupos – variáveis sociodemográficas

Faixa etária (em anos)							
Fator	Grupo — (N)	Média	<i>dp</i>	<i>t</i>	<i>gl</i>	<i>p</i>	<i>d</i>
AC	até 40 — (24)	2,97	0,72	-0,04	39	0,971	0,01
	41 + — (17)	2,98	0,69				
PB	até 40 — (24)	2,70	0,46	0,28	39	0,781	0,09
	41 + — (17)	2,66	0,47				
PMI	até 40 — (24)	2,67	0,39	0,90	39	0,374	0,28
	41 + — (17)	2,55	0,44				
TR	até 40 — (24)	2,71	0,38	1,17	39	0,248	0,38
	41 + — (17)	2,59	0,28				
LA_Geral	até 40 — (24)	2,76	0,41	0,56	39	0,578	0,18
	41 + — (17)	2,69	0,36				
Sexo (M ou F)							
Fator	Grupo - (N)	Média	<i>dp</i>	<i>t</i>	<i>gl</i>	<i>p</i>	<i>d</i>
AC	M — (24)	2,99	0,72	0,11	39	0,911	0,04
	F — (17)	2,96	0,70				
PB	M — (24)	2,65	0,45	-0,54	39	0,591	0,17
	F — (17)	2,73	0,47				
PMI	M — (24)	2,63	0,46	0,26	39	0,799	0,08
	F — (17)	2,60	0,35				
TR	M — (24)	2,64	0,34	-0,42	39	0,674	0,13
	F — (17)	2,69	0,35				
LA_Geral	M — (24)	2,73	0,40	-0,13	39	0,870	0,04
	F — (17)	2,74	0,38				
Função de liderança (ocupando ou não)							
Fator	Grupo - (N)	Média	<i>dp</i>	<i>t</i>	<i>gl</i>	<i>p</i>	<i>d</i>
AC	N — (18)	3,22	0,68	2,07	39	0,045	0,65
	S — (23)	2,78	0,67				
PB	N — (18)	2,78	0,29	1,27	35	0,246	0,38
	S — (23)	2,61	0,55				
PMI	N — (18)	2,69	0,43	0,92	39	0,361	0,29
	S — (23)	2,57	0,40				
TR	N — (18)	2,78	0,27	2,00	39	0,053	0,64
	S — (23)	2,57	0,37				
LA_Geral	N — (18)	2,87	0,34	2,00	39	0,053	0,64
	S — (23)	2,63	0,40				
Tamanho da empresa (por número de funcionários)							
Fator	Grupo - (N)	Média	<i>dp</i>	<i>t</i>	<i>gl</i>	<i>p</i>	<i>d</i>
AC	até 100 — (25)	2,92	0,63	-0,63	39	0,532	0,20
	101 + — (16)	3,06	0,82				
PB	até 100 — (25)	2,67	0,47	-0,19	39	0,851	0,06
	101 + — (16)	2,70	0,45				
PMI	até 100 — (25)	2,63	0,38	0,30	39	0,768	0,09
	101 + — (16)	2,59	0,47				
TR	até 100 — (25)	2,65	0,30	-0,25	39	0,807	0,08
	101 + — (16)	2,68	0,40				
LA_Geral	até 100 — (25)	2,72	0,34	-0,32	39	0,753	0,10
	101 + — (16)	2,76	0,45				

encontrado por Catano et al. (2012) em seu próprio experimento ($r = 0,82$; sendo o outro achado de $r = 0,66$).

Neste sentido, características distintas dos experimentos podem ser promotoras das diferenças obtidas. A primeira se refere ao período entre as coletas, que no estudo de Catano et al. (2012) foi de duas semanas e neste estudo, de mais de 12 meses. Trata-se de um período suficientemente amplo para que o nível de LA dos indivíduos, não sendo um traço, mas sim um atributo que pode ou não ser nutrido, de acordo com a perspectiva desenvolvimental

(Northhouse, 2018), possa sofrer modificações.

A segunda diz respeito à forma de pontuação dos testes aplicados que, reconhecidamente, influencia a confiabilidade dos TJS (De Leng et al., 2017; Tiffin et al., 2019). O teste utilizado por Catano et al. (2012) tem seus escores baseados em uma única seleção de opção (o que o respondente faria). No caso da EALA, os respondentes selecionaram tanto o que provavelmente fariam (EALA+), quanto o que não fariam (EALA-) e ambas as decisões geraram uma única pontuação, expandindo a multidimensionalidade

dade do item explanada por Campos e Rueda (2020) também para as respostas, conforme Martin-Raugh et al. (2018).

No caso do fator PMI, que obteve o menor valor de r , é necessário considerar também o desafio inerente a essa medida, uma vez que a teoria da LA não especifica estritamente os valores que um líder autêntico deve apresentar, mas apenas que ele deve tomar decisões congruentes com seus valores pessoais internalizados e que esses são sempre de cunho positivo (Avolio & Gardner, 2005; Gardner et al., 2005; Walumbwa et al., 2008). Para Weiss et al. (2018) esse aspecto é crítico com relação aos resultados que podem ser obtidos por meio de avaliação em modo de heterorrelato (liderados não teriam como assegurar a coesão de valores do líder). Esse aspecto parece constituir, também, uma das dificuldades para a avaliação do líder de forma objetiva e direta, diante da amplitude da teoria.

Certamente a teoria da LA permite entender que ética, honestidade e integridade são alguns dos valores a serem apresentados por um líder autêntico (Avolio & Gardner, 2005). Seria contributivo, entretanto, evidenciar não apenas que outros valores devem ser considerados, como também como deve agir um líder autêntico em situações de conflito entre dois ou mais de seus valores pessoais.

A habilidade de lidar com as situações de conflito é aderente (se não parte integrante) da proposta feita por Johnsen (2018) para uma nova conceituação da LA. Para esse autor, a autenticidade de um líder não deve ser definida pelo seu comprometimento com seus valores pessoais, mas sim por sua capacidade de refletir sobre quanto esse comprometimento pode aproximá-lo ou afastá-lo de uma decisão ética. Assim, o desenvolvimento de novos itens para a EALA agregando situações que possam avaliar essa capacidade reflexiva por parte do líder e analisar as diferenças daí advindas seria contributivo para o avanço do campo de conhecimento.

Além de considerar as correlações encontradas neste estudo, importa também considerar que as diferenças de médias obtidas para os dois momentos de coleta não se mostraram significativas e que os tamanhos de efeito encontrados pelo d de Cohen, variando entre 0,01 e 0,41, podem ser interpretados como insignificantes ou pequenos de acordo com Cohen (1988) e com Espírito-Santo e Daniel (2015). Desta forma, a despeito das considerações teóricas realizadas, que demandam ainda um amplo debate, considera-se que os resultados obtidos pela investigação de teste-reteste são bons indicadores iniciais da confiabilidade da EALA. Considera-se que novos experimentos de teste-reteste, com períodos menores entre as coletas, serão bem-vindos para que esses resultados possam ser corroborados.

Ainda, poderão também ser estudadas formas alternativas de pontuação da EALA. Isso possibilitará a comparação de evidências de validade e de confiabilidade, a exemplo do que foi realizado por De Leng et al. (2017) e de outros estudos mencionados por Catano et al. (2012). Tais procedimentos ajudarão, também, a corroborar se as qualidades psicométricas do instrumento demonstram viabilidade para sua aplicação na prática organizacional. Nesse sentido, novas pesquisas realizadas no campo da prática organizacional devem ser promovidas, visto que podem ampliar a compreensão sobre a adequação do instrumento às aplicações para as quais foi proposto (i.e., seleção e desenvolvimento de líderes).

Com relação ao segundo objetivo deste artigo, analisar diferenças entre grupos, à exceção do fator autoconsciência no caso de indivíduos exercendo ou não funções de liderança, não foram encontradas diferenças estatisticamente significativas. Importa discutir esses aspectos à luz de conceitos relativos à psicometria e à teoria da LA, incluindo as críticas ao modelo vigente. No que diz respeito aos aspectos psicométricos, importa considerar o tama-

nho da amostra que participou deste estudo ($N = 41$), que viabilizou a distinção entre grupos com um número ainda menor de participantes. Desta forma, cabe recordar que, sendo o tamanho da amostra o elemento mais determinante no cálculo da significância de p , conforme as considerações de literatura recuperadas por Espírito-Santo e Daniel (2015), os resultados obtidos neste estudo demandam que se dê atenção não apenas ao nível de significância estatística, mas também aos tamanhos de efeito encontrados, uma vez que a significância prática pode ampliar o entendimento do contexto.

No que diz respeito à investigação das diferenças entre os grupos caracterizados por sexo, a inexistência de significância estatística e os tamanhos de efeito insignificantes indicam que as críticas à teoria da LA apresentadas por Hopkins e O'Neil (2015) não se confirmam na prática. Empiricamente, as mulheres não demonstraram menor possibilidade de assumir a autenticidade como estilo de liderança. Esses resultados corroboram os achados de Cervo et al. (2018) que, utilizando o ALQ, obtiveram 277 líderes brasileiros (52,7% mulheres) avaliados por 310 liderados (62,6% mulheres) e não encontraram diferenças estatisticamente significativas entre os grupos formados por sexo. Campos e Rueda (2019) também encontraram resultados semelhantes, por meio de um estudo de cunho qualitativo que analisou o discurso de dez líderes contemporâneos brasileiros, dentre os quais três eram mulheres. Desta forma, esses achados são indícios de que não há necessidade de se complementar ou modificar a teoria da LA nesse campo e são de interesse para organizações que têm buscado ampliar a diversidade e a justa oferta de oportunidades para profissionais independente de gênero.

No caso dos grupos formados pela distinção de faixa etária, cabe recordar que a teoria da LA sob a perspectiva desenvolvimental indica que esse estilo se desenvolve ao longo da vida e que esse desenvolvimento tanto pode se dar devido a experiências vivenciadas pelo líder, quanto por meio de intervenções específicas para isso, o que Northhouse (2018) chamou de nutrir o estilo. Tais pressupostos teóricos poderiam conduzir à hipótese de que quanto maior a idade do líder, o que pode implicar em um número maior de experiências vividas por ele, maior seria a possibilidade de esse líder se mostrar um líder autêntico.

Nesse sentido, apesar de não terem sido encontradas diferenças estatisticamente significativas com base na faixa etária, para dois dos fatores (PMI, $d = 0,28$ e TR, $d = 0,38$) foram encontrados pequenos tamanhos de efeito (Cohen, 1988; Espírito-Santo & Daniel, 2015) que indicam o contrário. Isto é, os indivíduos com menor faixa etária obtiveram melhores pontuações para esses fatores e o efeito dessas diferenças não foi insignificante. Esses achados promovem algumas questões a serem respondidas por novas pesquisas que possam ser realizadas com amostras maiores, unidas a uma investigação sobre a cultura organizacional. Uma dessas questões é: as organizações atuais constituem ambientes que, por seu contexto altamente competitivo e voltado ao alcance de resultados financeiros cada vez mais desafiadores, podem promover distanciamento dos indivíduos que ali trabalham com relação a seus valores núcleo (fator PMI) e à sua capacidade de se posicionarem francamente em termos de crenças, verdades e perspectivas (fator TR)?

De certa forma, é isso que Alvesson e Einola (2019) criticaram na teoria da LA, posicionando que ela carrega uma visão fantasiosa e distanciada do real ambiente organizacional. Esses autores afirmaram que “a comunidade de pesquisa precisa começar a considerar as contradições e dilemas da vida real, em vez de ser enganada pela ideologia”. Entretanto, essa visão não invalida a teoria da LA por completo. Apenas demanda que ela seja aprofundada, ampliando conceitos ainda não bem delineados. A respeito

do fator PMI, por exemplo, os autores questionaram para que direção “moral” de “fazer o bem” o líder deve se voltar: acionistas, subordinados, colegas de trabalho, planeta, contribuintes, grupos de interesse?

De fato, essa é uma das complexidades do comportamento organizacional que é normalmente respondida por sua cultura. Os resultados obtidos neste estudo para o fator PMI são indícios de que essa complexidade é real na prática. Entretanto, o formato de TJS escolhido para a EALA pode ser uma forma de resolver essa questão no âmbito de medidas, pois é possível desenvolver cenários que avaliem e estimem a hierarquia de prioridades morais que o líder pode assumir em seu processo de tomada de decisão. O mesmo pode ser considerado a respeito do fator TR. Para Alvesson e Einola (2019), “muita” autenticidade no local de trabalho pode prejudicar o clima e propiciar o surgimento de conflitos. Mas o que caracterizaria “muita”? Mais uma vez, um melhor delineamento para o fator pode ser apresentado pela teoria e o formato de teste de julgamento situacional pode ser útil para a compreensão dessa hierarquia, uma vez que não se pretende encontrar um líder que seja totalmente autêntico, mas sim que se possa diferenciar se ele é mais ou se é menos autêntico (Gardner et al., 2005).

No intuito de avaliar se um líder é mais ou menos autêntico, os resultados obtidos neste estudo para as diferenças entre os grupos caracterizados por estar ou não exercendo funções de liderança apresentaram indícios de que indivíduos que se encontravam efetivamente responsáveis por liderar outros indivíduos se mostraram menos autênticos do que os que não estavam exercendo essas funções. Esses resultados merecem atenção e promovem um conjunto de indagações que pode endereçar pesquisas futuras. Algumas dessas questões são: 1) a teoria da LA define um “comportamento de líder” que iria se inviabilizando na medida em que o “papel de líder” no real ambiente organizacional vai sendo introjetado pelo indivíduo?; 2) o exercício do “papel de líder” promovido por escolas de liderança e pela prática organizacional é incongruente com o exercício da autoconsciência (conhecer-se e também ao impacto que provoca sobre outros), do processamento balanceado (ser capaz de ouvir e refletir sobre ideias diferentes das suas) da perspectiva moral internalizada (fidelidade aos valores pessoais) e da transparência (*self* autêntico)?; 3) em que medida uma potencial incongruência entre esses aspectos pode promover sofrimento (ou bem-estar) para líderes e liderados? Estudos longitudinais aplicando a EALA com grupos que passem por intervenções para desenvolvimento de liderança baseados na LA e no *status quo* podem ser delineados e conduzidos com o objetivo de responder a essas questões.

Tais investigações poderão colaborar, inclusive, para atenuar limitações deste estudo. Algumas delas são: a pequena participação da amostra convidada para o reteste, o período entre as coletas e a ausência de controle sobre eventuais processos de desenvolvimento de liderança ao longo do período entre as coletas. Entretanto, as limitações não impediram que o estudo cumprisse seus objetivos de investigação. Os índices iniciais de confiabilidade se mostraram promissores para a continuidade de estudos com a EALA, possibilitando, inclusive uma apreciação teórica da LA. Do mesmo modo, os resultados da análise de diferenças entre grupos apontaram caminhos para novas investigações que poderão contribuir com os campos empírico e teórico.

Referências

Algera, P. M., & Lips-Wiersma, M. (2012). Radical authentic leadership: Co-creating the conditions under which all members of the organization can be authentic. *The Leadership Quarterly*, 23(1), 118-131. <https://doi.org/10.1016/j.leaqua.2011.11.010>

- Alvesson, M., & Einola, K. (2019). Warning for excessive positivity: Authentic leadership and other traps in leadership studies. *The Leadership Quarterly*, 30(4), 383-395. <https://doi.org/10.1016/j.leaqua.2019.04.001>
- Ambiel, R. A. M., Campos, M. I., Alves, B. D. P., & Silva, C. P. (2015). Fundamentos e aplicabilidade dos Testes de Julgamento Situacional no contexto da Psicologia Organizacional. *Revista Psicologia Organizações e Trabalho*, 15(3), 298-308. <https://doi.org/10.17652/rpot/2015.3.529>
- Antonakis, J. (2017). Charisma and the “New Leadership”. Em J. Antonakis & D. V. Day (eds) (2017), *The Nature of Leadership*. Thousand Oaks (pp. 56-81), CA: Sage.
- Avolio, B. J., & Gardner, W. L. (2005). Authentic leadership development: Getting to the root of positive forms of leadership. *The Leadership Quarterly*, 16(3), 315-338. <https://doi.org/10.1016/j.leaqua.2005.03.001>
- Avolio, B. J., Gardner, W. L., Walumbwa, F. O., Luthans, F., & May, D. R. (2004). Unlocking the mask: A look at the process by which authentic leaders impact follower attitudes and behaviors. *The Leadership Quarterly*, 15(6), 801-823. <https://doi.org/10.1016/j.leaqua.2004.09.003>
- Avolio, B. J., & Walumbwa, F. O. (2014). Authentic leadership theory, research and practice: Steps taken and steps that remain. Em D. V. Day (Ed.), *The Oxford handbook of leadership and organizations* (pp. 331-356). New York, NY: Oxford University Press.
- Avolio, B. J., Wernsing, T., & Gardner, W. L. (2018). Revisiting the Development and Validation of the Authentic Leadership Questionnaire: Analytical Clarifications. *Journal of Management*, 44(2), 399-411. <https://doi.org/10.1177/0149206317739960>
- Banks, G. C., Gooty, J., Ross, R. L., Williams, C. E., & Harrington, N. T. (2018). Construct redundancy in leader behaviors: A review and agenda for the future. *The Leadership Quarterly*, 29(1), 236-251. <https://doi.org/10.1016/j.leaqua.2017.12.005>
- Banks, G. C., McCauley, K. D., Gardner, W. L., & Guler, C. E. (2016). A meta-analytic review of authentic and transformational leadership: A test for redundancy. *The Leadership Quarterly*, 27(4), 634-652. <https://doi.org/10.1016/j.leaqua.2016.02.006>
- Baron, L. (2016). Authentic leadership and mindfulness development through action learning. *Journal of Managerial Psychology*, 31(1), 296-311. <https://doi.org/10.1108/JMP-04-2014-0135>
- Campos, M. I., & Rueda, F. J. M. (2020). Authentic leadership: Development and initial validation of a situational judgment test. *Revista Psicologia: Organizações e Trabalho*, 20(2), 1047-1056. <https://doi.org/10.17652/rpot/2020.2.18100>
- Campos, M. I., & Rueda, F. J. M. (2019). Authentic leadership: A theoretical thematic analysis of the contemporary Brazilian leader's speech. *Paidéia (Ribeirão Preto)*, 29, e2924. <https://doi.org/10.1590/1982-4327e2924>
- Campos, M. I., & Rueda, F. J. M. (2018a). *Authentic Leadership Measures: a Literature Review*. Manuscrito submetido para publicação.
- Campos, M. I., & Rueda, F. J. M. (2018b). Evolução do construto liderança autêntica: uma revisão de literatura. *Revista Psicologia Organizações e Trabalho*, 18(1), 291-298. <https://doi.org/10.17652/rpot/2018.1.13473>
- Catano, V. M., Brochu, A., & Lamerson, C. D. (2012). Assessing the reliability of situational judgment tests used in high-stakes situations. *International Journal of Selection and Assessment*, 20(3), 333-346. <https://doi.org/10.1111/j.1468-2389.2012.00604.x>
- Cervo, C. S., Natividade, J. C., Mónico, L. D. S. M., Pais, L., dos Santos, N. R., & Hutz, C. S. (2018). Modelo de Liderança Autêntica: concepção teórica e evidências de validade do Authentic Leadership Questionnaire (ALQ) para o Brasil. *Psicologica*, 61(2), 7-29. https://doi.org/10.14195/1647-8606_61-2_1
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd. ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Corstjens, J., Lievens, F., & Krumm, S. (2017). Situational judgment tests for selection. Em H. W. Goldstein, E. D. Pulakos, J. Passmore, & C. Semedo (Eds.), *The Wiley Blackwell handbook of the psychology of recruitment, selection and employee retention* (pp. 226-246). Oxford, UK: Wiley. <https://doi.org/10.1002/9781118972472.ch11>
- Dancey, C. P., & Reidy, J. (2006). *Estatística sem matemática para psicologia usando SPSS para Windows* (3ª Ed.). Porto Alegre, RS: Artmed.
- De Leng, W. E., Stegers-Jager, K. M., Husbands, A., Dowell, J. S., Born, M. P., & Themmen, A. P. N. (2017). Scoring method of a Situational Judgment Test: influence on internal consistency reliability, adverse impact and correlation with personality?. *Advances in Health Sciences Education*, 22(2), 243-265. <https://doi.org/10.1007/s10459-016-9720-7>
- Espírito-Santo, H., & Daniel, F. (2015). Calcular e apresentar tamanhos do efeito em trabalhos científicos (1): As limitações do $p < 0,05$ na análise de diferenças de médias de dois grupos. *Revista Portuguesa de Investigação Comportamental e Social*, 1(1), 3-16. <https://doi.org/10.7342/ismt.rpics.2015.1.1.14>
- Ford, J., & Harding, N. (2011). The impossibility of the ‘true self’ of authentic leadership. *Leadership*, 7(4), 463-479. <https://doi.org/10.1177/1742715011416894>

- Fusco, T. (2018). *An Evidence-based Approach to Authentic Leadership Development*. London, UK: Routledge. <https://doi.org/10.4324/9781315187990>
- Gardner, W. L., Avolio, B. J., Luthans, F., May, D. R., & Walumbwa, F. O. (2005). "Can you see the real me?" A self-based model of authentic leader and follower development. *The Leadership Quarterly*, 16(3), 343-372. <https://doi.org/10.1016/j.leaqua.2005.03.003>
- Goss, B. D., Ryan, A. T., Waring, J., Judd, T., Chiavaroli, N. G., O'Brien, R. C., Trumble, S. C., & McColl, G. J. (2017). Beyond selection: the use of situational judgement tests in the teaching and assessment of professionalism. *Academic Medicine*, 92(6), 780-784. <https://doi.org/10.1097/ACM.0000000000001591>
- Hopkins, M. M., & O'Neil, D. A. (2015). Authentic leadership: application to women leaders. *Frontiers in Psychology*, 6, 1-5. <https://doi.org/10.3389/fpsyg.2015.00959>
- Johnsen, C. G. (2018). Authenticating the leader: Why Bill George believes that a moral compass would have kept Jeffrey Skilling out of jail. *Journal of Business Ethics*, 147(1), 53-63. <https://doi.org/10.1007/s10551-015-2968-7>
- Kaminski, K., Felfe, J., Schäpers, P., & Krumm, S. (2019). A closer look at response options: Is judgment in situational judgment tests a function of the desirability of response options?. *International Journal of Selection and Assessment*, 27(1), 72-82. <https://doi.org/10.1111/ijsa.12233>
- Kempster, S., Iszatt-White, M., & Brown, M. (2018). Authenticity in leadership: Reframing relational transparency through the lens of emotional labour. *Leadership*, 15(3), 319-338. <https://doi.org/10.1177/1742715017746788>
- Levesque-Côté, J., Fernet, C., Austin, S., & Morin, A. J. (2017). New wine in a new bottle: Refining the assessment of authentic leadership using exploratory structural equation modeling (ESEM). *Journal of Business and Psychology*, 1-18. <https://doi.org/10.1007/s10869-017-9512-y>
- Martin-Raugh, M. P., Anguiano-Carrasco, C., Jackson, T., Brennehan, M. W., Carney, L., Barnwell, P., & Kochert, J. (2018). Effects of Situational Judgment Test Format on Reliability and Validity. *International Journal of Testing*, 18(2), 135-154. <https://doi.org/10.1080/15305058.2018.1428981>
- Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*, 75(6), 640-647. <https://doi.org/10.1037/0021-9010.75.6.640>
- Neider, L. L., & Schriesheim, C. A. (2011). The authentic leadership inventory (ALI): Development and empirical tests. *The Leadership Quarterly*, 22(6), 1146-1164. <https://doi.org/10.1016/j.leaqua.2011.09.008>
- Northouse, P. G. (2018). *Leadership: Theory and practice* (8th ed.). Los Angeles: Sage.
- Sidani, Y. M., & Rowe, W. G. (2018). A reconceptualization of authentic leadership: Leader legitimation via follower-centered assessment of the moral dimension. *The Leadership Quarterly*, 29(6), 623-636. <https://doi.org/10.1016/j.leaqua.2018.04.005>
- Sorrel, M. A., Olea, J., Abad, F. J., de la Torre, J., Aguado, D., & Lievens, F. (2016). Validity and reliability of situational judgement test scores: A new approach based on cognitive diagnosis models. *Organizational Research Methods*, 19(3), 506-532. <https://doi.org/10.1177/1094428116630065>
- Tiffin, P. A., Paton, L. W., O'Mara, D., MacCann, C., Lang, J. W., & Lievens, F. (2019). Situational judgement tests for selection: Traditional vs construct-driven approaches. *Medical Education*, 54, 105-115. <https://doi.org/10.1111/medu.14011>
- Walumbwa, F. O., Avolio, B. J., Gardner, W. L., Wernsing, T. S., & Peterson, S. J. (2008). Authentic leadership: Development and validation of a theory-based measure. *Journal of Management*, 34(1), 89-126. <https://doi.org/10.1177/0149206307308913>
- Weiss, M., Razinskas, S., Backmann, J., & Hoegl, M. (2018). Authentic leadership and leaders' mental well-being: An experience sampling study. *The Leadership Quarterly*, 29(2), 309-321. <https://doi.org/10.1016/j.leaqua.2017.05.007>

Informações sobre os autores:

Maria Isabel de Campos

Universidade São Francisco, Campus Swift.
Rua Waldemar César da Silveira, 105. Jd. Cura D'Ars.
13045-510 Campinas, SP, Brasil
E-mail: isabel.playit@gmail.com

Fabián Javier Marín Rueda

E-mail: marinfabian@gmail.com