



Universidade de Brasília
Instituto de Psicologia
Departamento de Psicologia Social e do Trabalho
Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações

Thiago Augusto Costa de Olival

A CRENÇA EM CIÊNCIA, RACIONALIDADE E TRAÇOS DE PERSONALIDADE E SUAS
RELAÇÕES COM O JULGAMENTO E DECISÃO EM AMBIENTE ORGANIZACIONAL

Brasília
2021

THIAGO AUGUSTO COSTA DE OLIVAL

A CRENÇA EM CIÊNCIA, RACIONALIDADE E TRAÇOS DE PERSONALIDADE E SUAS
RELAÇÕES COM O JULGAMENTO E DECISÃO EM AMBIENTE ORGANIZACIONAL

SCIENCE'S BELIEF, RATIONALITY AND PERSONALITY TRAITS AND THEIR
RELATIONSHIPS WITH JUDGEMENT AND DECISION MAKING IN ORGANIZATIONAL
ENVIRONMENT

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações, do Instituto de Psicologia da Universidade de Brasília, como parte dos requisitos para obtenção do título de Mestre em Psicologia Social, do Trabalho e das Organizações.

Orientadora: Prof.^a Dra. Elaine Rabelo Neiva

Brasília,
Maio de 2021

THIAGO AUGUSTO COSTA DE OLIVAL

A CRENÇA EM CIÊNCIA, RACIONALIDADE E TRAÇOS DE PERSONALIDADE E SUAS
RELAÇÕES COM O JULGAMENTO E DECISÃO EM AMBIENTE ORGANIZACIONAL

SCIENCE'S BELIEF, RATIONALITY AND PERSONALITY TRAITS AND THEIR
RELATIONSHIPS WITH JUDGEMENT AND DECISION MAKING IN ORGANIZATIONAL
ENVIRONMENT

Dissertação de mestrado apresentada ao
Programa de Pós-graduação em Psicologia
Social, do Trabalho e das Organizações, do
Instituto de Psicologia da Universidade de
Brasília, como parte dos requisitos para
obtenção do título de Mestre em Psicologia
Social, do Trabalho e das Organizações.

Esta dissertação foi avaliada pela seguinte comissão examinadora:

Prof.^a Dra. Elaine Rabelo Neiva (Presidente)
Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações
Universidade de Brasília (UnB)

Prof. Dr. Nelson Hauck Filho (Membro titular)
Programa de Pós-graduação em Psicologia
Universidade São Francisco (USF)

Prof. Dr. Luís Gustavo do Amaral Vinha (Membro titular)
Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações
Universidade de Brasília (UnB)

Prof.^a Dr. Josemberg Moura de Andrade (Membro suplente)
Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações
Universidade de Brasília (UnB)

Agradecimentos

Aos meus pais e avós, pela dedicação irrestrita e investimento em minha educação e criação, aos quais sou mais do que grato.

A minha esposa Fernanda, companheira de jornada e incentivadora, meu amor e mãe da princesa Monica. Sem elas, não teria conseguido.

A professora Elaine Neiva pela disponibilidade e pelos “nãos”, tão imprescindíveis para o aprendizado. Mais do que procedimentos estatísticos complexos, ensinou-me o que é ciência na prática.

Ao professor Itiro, que apontou caminhos importantes em momentos de indecisão.

Ao professor André Rabelo, pelas breves, mas assertivas discussões.

Ao professor Ronaldo Pilati, pela abertura e interesse em partilhar seus conhecimentos.

Ao professor Nelson Hauck, que sempre foi muito solícito e interessado no projeto, cuja parceria foi fundamental para a conclusão desta jornada.

Aos colegas do grupo de pesquisa Inovare UNB pela constante troca de informações e idéias, que com as discussões contribuíram para a elaboração e melhoria deste trabalho.

Aos demais alunos da graduação, mestrandos e doutorandos com os quais tive alguma interlocução da mais rápida até a mais costumeira. Mais do que discutir teorias, discutimos a vida de diferentes maneiras e perspectivas.

Aos colegas do Banco do Brasil, que apoiaram e possibilitaram de alguma forma que esta pesquisa fosse concluída com sucesso.

Resumo

O processo decisório organizacional é complexo e possui considerável influência de variáveis internas (estilo cognitivo, crenças, personalidade e conhecimentos) e externas (cenário, quantidade e qualidade das informações), de maneira que carecem de estudos sobre julgamento e decisão em contexto aplicado a realidade organizacional e que possam diferenciar características relacionadas a boas decisões envolvendo gestão de pessoas e gestão financeira, além de explorar psicometricamente as características de variáveis requeridas para as decisões. A dissertação está organizada em 3 manuscritos considerando a busca por evidência de validade em dois instrumentos adaptados ao contexto brasileiro e uma criação inédita de teste situacional de julgamento e decisão aplicado a realidade organizacional, onde também são observados indícios de validade. O manuscrito 1 apresenta o processo de adaptação cultural do instrumento *Cognitive Reflection Test* (CRT) para o contexto brasileiro, além de indícios de validade decorrentes do processo de tradução e julgamento por juízes, análise fatorial exploratória e confirmatória, além da técnica TRI (teoria de resposta ao item) do instrumento. No segundo manuscrito, são apresentados os indícios de validade do instrumento *Believe in Science Scale* (BISS) decorrentes do processo de adaptação cultural ao contexto brasileiro e da realização de análise fatorial exploratória e confirmatória, da análise da correlação do instrumento com outros construtos e fatores, além da comparação entre as médias de crença na ciência entre grupos religiosos. No 3 manuscrito, são relatados os procedimentos de criação e observação de indícios de validade de testes situacionais de julgamento e decisão em cenários onde são manipuladas situações envolvendo finanças e gestão de pessoas, por meio da análise de classes latentes, comparação entre grupos e correlação entre as decisões e outros construtos e fatores. Em conjunto, espera-se que os manuscritos possam contribuir para novos estudos envolvendo racionalidade, crença em ciência e julgamento e decisão em contexto aplicado as organizações. Ao final, são discutidas as limitações e sugestão para trabalhos futuros.

Palavras-chave: Racionalidade, crença em ciência, julgamento e decisão

Abstract

The organizational decision-making process is complex and has considerable influence of internal variables (cognitive style, beliefs, personality and knowledge) and external ones (scenario, quantity and quality of information), so that studies on judgment and decision in context applied to organizational reality and that can differentiate characteristics related to good decisions involving people management and financial management, in addition to psychometrically exploring the characteristics of variables required for decisions. The dissertation is organized in 3 manuscripts considering the search for evidence of validity in two instruments adapted to the Brazilian context and an unprecedented creation of a situational test of judgment and decision applied to organizational reality, where signs of validity are also observed. Manuscript 1 presents the process of cultural adaptation of the Cognitive Reflection Test (CRT) for the Brazilian context, in addition to evidence of validity resulting from the process of translation and judgment by judges, exploratory and confirmatory factor analysis, in addition to the realization of IRT (Item Response Theory) of the instrument. In the second manuscript, evidence of the validity of the Believe in Science Scale (BISS) instrument resulting from the process of cultural adaptation to the Brazilian context and the carrying out of exploratory and confirmatory factor analysis, from the analysis of the correlation of the instrument with other constructs and factors, are presented, in addition to the comparison between the means of science belief between religious groups. In the 3rd manuscript, the procedures for creating and observing evidence of validity of situational tests of judgment and decision in scenarios where situations involving finance and people management are manipulated, through the analysis of latent classes, comparison between groups and correlation between decisions and other constructs and factors. Together, it is expected that the manuscripts can contribute to new studies involving rationality, belief in science and judgment and decision in a context applied to organizations. At the end, limitations and suggestions for future work are discussed.

Keywords: Rationality, Believ in Science, Judgement and decision-making

Sumário

Sumário

Lista de Figuras.....	9
Lista de Tabelas.....	10
Apresentação.....	11
Referências.....	14
Manuscrito 1 - Rápido ou Devagar? – Evidências de Validade do Teste de Reflexão Cognitiva (CRT) para o Contexto Brasileiro	16
Resumo	17
Abstract	18
Introdução.....	19
O Sistema Dual, Suas Possíveis Métricas e Relações.....	19
Mensuração da Reflexão Cognitiva no Contexto Brasileiro.....	22
Método	22
Adaptação do Teste ao Contexto do Brasil.....	22
Tradução do Instrumento.....	22
Síntese do Instrumento, Avaliação Por Especialistas e Pelo Público Alvo	23
Tradução Reversa	24
Participantes.....	24
Análise dos Dados.....	25
Resultados.....	25
Análise Fatorial.....	26
Análise Fatorial Exploratória	26
TRI (Teoria de Resposta ao Item).....	28
Comparação Entre Resultados Brasileiros e Estudos de Outros Países.....	30
Discussão	31
Referências	34
Apêndice A - Teste de Reflexão Cognitiva (Cognitive Reflexion Test – CRT).....	42
Manuscrito 2 - Uma Vela na Escuridão: Evidências de Validade da Escala de Crença em Ciência (BISS) Para o Contexto Brasileiro.	43
Resumo	44
Abstract	45
Introdução.....	46
Referencial Teórico.....	47
A Crença na Ciência.....	47
Letramento Científico Brasileiro	48

O Brasileiro e a Crença na Ciência	48
Método	49
Adaptação ao Contexto Brasileiro.....	49
Variáveis Presentes no Manuscrito.....	51
Participantes e Pressupostos Estatísticos.....	51
Resultados.....	53
Estudo 1 - Análise Fatorial Exploratória	53
Estudo 2 – Crença na Ciência e Suas Correlações	55
Estudo 3 – Grupos Religiosos e a Crença na Ciência.....	56
Discussão	57
Referências	61
Apêndice B – Escala de Crença em Ciência.....	70
Manuscrito 3 – Uma Decisão com Classe: Análise de Classes Latentes e Índicios de Validade em Testes Situacionais de Julgamento	71
Resumo	72
Abstract	73
Introdução.....	74
Referencial Teórico.....	75
A Utilização de Testes Situacionais de Julgamento	75
Testes Situacionais Envolvendo Decisões Financeiras.....	76
Bases Para a Decisão	77
Índicios de Validade em Testes Situacionais	78
Método	80
Construção do Teste de Decisão Organizacional Aplicado a Finanças e a Gestão de Pessoas.....	80
Descrição dos Cenários.....	81
Participantes.....	84
Procedimento de Coleta de Dados	85
Procedimento de Análise de Dados	85
Resultados.....	87
Análise das Classes Latentes	87
Comparação Entre Grupos e Classes.....	89
Correlação Entre Padrão de Respostas e Variáveis	89
Discussão	90
Referências	93
Apêndice C – Testes Situacionais de Julgamento e Decisão.....	108
Considerações Finais.....	113

Lista de Figuras

Figura 1.1 – Comparativo entre padrão CCI (Curva Característica do Item) e CII (Curva de Informação do Item)	30
Figura 2.1 Comparativo da distribuição das classes entre grupos	88

Lista de Tabelas

Tabela 1.1 Caracterização dos participantes da pesquisa.....	25
Tabela 1.2 Relação dos índices da análise	27
Tabela 1.3 Índices dos itens do modelo	29
Tabela 1.4 Comparativo entre estudos.....	30
Tabela 2.1 Comparação da versão final da pesquisa com outra versão já utilizada	50
Tabela 2.2 Caracterização dos participantes da pesquisa.....	52
Tabela 2.3 Relação dos demais índices da análise	54
Tabela 2.4 Correlação de Spearman da crença com outros fatores	55
Tabela 2.5 Comparação entre grupos via teste Tukey	56
Tabela 3.1 Resumo das questões e induções inseridas nos cenários.....	84
Tabela 3.2 Caracterização dos participantes da pesquisa.....	85
Tabela 3.3 Comparação entre os modelos da pesquisa.....	87
Tabela 3.4 Correlação entre cenários e variáveis do modelo	90

Apresentação

O processo de julgamento e decisão empresarial é constante e de extrema importância para o diferenciar sucesso de fracasso, na medida em que decisões ruins podem ocasionar prejuízos tamanhos que podem culminar no encerramento das atividades de uma organização. O entendimento das variáveis que podem influenciar o processo decisório e limitar a racionalidade (Simon, 1955) tem sido contruído na história contemporânea e a questão psicológica começou a ganhar maior notoriedade com estudos de Amos Tversky e Daniel Kahneman em meados da década de 70. Esses trabalhos adicionaram ponderações sobre as evidências de que os indivíduos não tomam decisões racionais na mesma proporção e direção esperadas e que o processo decisório possui muitos vieses cognitivos e simplificações chamadas heurísticas (Tversky & Kahneman, 1974).

Tais simplificações estariam relacionadas a intuição enquanto que elocubrações mais elaboradas estariam relacionadas a racionalidade. Porém, de acordo com Pennycook et al. (2015), nem sempre uma resposta tida como racional necessariamente foi processada com alta carga cognitiva, mas poderia trata-se de uma intuição decorrente de alguém bem treinado ou experiente em determinado campo. Medir o construto racionalidade é complexo visto que a definição operacional transcende o objeto puro e simples da numerácia, e a medida de racionalidade possui muitas formas. Uma das formas mais populares observadas dentro e fora de pesquisas acadêmicas é decorrente do instrumento *Cognitive Reflection Test* (CRT) construído por Frederick (2005), cuja proposta é medir o fenômeno com um instrumento que contém 3 itens de avaliação lógico-matemática em que são transcritos problemas envolvendo operações matemáticas simples e que, em um primeiro momento, podem induzir o indivíduo a uma resposta equivocada, tida como intuitiva, e que poderia ser suplantada por uma resposta correta ao se equacionar melhor o problema.

Conceitos relacionados a racionalidade têm sido amplamente debatidos em pesquisas acadêmicas e apresentados em livros cujo tema trata de termos modernos relacionados a área de finanças e investimentos (finanças comportamentais), a intersecção

de áreas da Economia e da Psicologia (chamadas popularmente de psicologia econômica ou economia comportamental). Apesar da crescente popularidade de temas relacionados a educação financeira, existe a preocupação de que instrumentos construídos com este fim estejam de fato medindo o que se propõem a medir.

Outro conceito derivado da racionalidade diz respeito a crença, e em especial a crença em ciência, cujas evidências apontam para a associação com o pensamento analítico (Rutjens et al, 2018; Stagnaro, 2019). Na tentativa de mensurar o fenômeno de crença na ciência, Faria et al, (2013) desenvolveram um instrumento cujos itens tratam de expor ao indivíduo algumas afirmações relacionadas a ciência, mas que não necessitariam obrigatoriamente de letramento científico, de forma que se avaliaria o grau de concordância acerca da frase exposta.

Crenças diversas coexistem em um mesmo indivíduo e muitas vezes são utilizadas de maneira convergente (Irvin et al, 2005) para justificar conclusões e reduzir possíveis dissonâncias, utilizando-se de informações seletivas que permitam esta maneira de pensamento duplo se alinhar. Então, a crença religiosa pode coexistir com a crença não religiosa, mesmo que tal fato possa parecer paradoxal a princípio. E quando se fala em crença na ciência em organizações, ainda pairam constantes dúvidas sobre qual seria o nível da crença ou talvez suas relações com o mundo do trabalho, visto os poucos estudos sobre o tema.

Nas organizações, os indivíduos deparam-se com cenários desafiadores em que múltiplas variáveis são consideradas para que haja uma decisão. Como produto deste processamento, existem decisões que podem ser mais racionais, sobre a ótica de majorar resultados financeiros, reduzir perdas e escolher investimentos mais atrativos quando o objetivo envolve finanças. Quando o objetivo envolve gestão de pessoas, decisões racionais poderiam estar relacionadas a evitar um viés ao selecionar um candidato, julgar um processo disciplinar ou realizar uma demissão de maneira imparcial e ponderar sobre características de um indivíduo de maneira mais objetiva.

Uma das formas de identificar padrões que classifiquem os indivíduos de acordo com a maneira que fazem escolhas abriria caminhos inúmeros para entender formas de seleção mais objetivas, alocação de pessoas de acordo com características requeridas para o trabalho e, por fim, melhorar o processo de decisão capacitando os envolvidos para tomar melhores decisões. Estudos envolvendo processo de julgamento e decisão em ambiente organizacional apresentam evidências de que algumas das decisões possuem grande carga subjetiva (Cohen, 2019). Identificar instrumentos que apresentem indícios de validade é o objetivo maior desta dissertação. Adicionalmente, são objetivos desta dissertação identificar indícios de validade tanto das variáveis que ainda não possuem tais indicadores para o contexto brasileiro, como também identificar tais indícios nos testes situacionais aqui desenvolvidos. Desta forma, são apresentados três manuscritos em que cada um dos principais temas expostos (racionalidade, crença em ciência e testes situacionais) são estudados de maneira mais específica.

No manuscrito 1, é apresentada uma versão adaptada do instrumento CRT para o contexto brasileiro, em que são identificados indícios de validade interna decorrentes do processo de adaptação (tradução, tradução reversa e avaliação por juízes), a partir da realização da análise fatorial exploratória e análise fatorial confirmatória, além da utilização da técnica chamada TRI (Teoria de Resposta ao Item), para que sejam mensuradas as características do item como dificuldade, discriminação e a curva de informações. Também são feitas comparações dos resultados encontrados no presente estudo com outros estudos internacionais.

No manuscrito 2, são apresentadas algumas das evidências de validade interna e externa do instrumento de crença em ciência (BISS) que perpassam pelo processo de tradução, tradução reversa, análise de juízes e comparação com outra versão do instrumento. No estudo, também estão presentes a análise fatorial exploratória e análise fatorial confirmatória, bem como um estudo das correlações da crença na ciência com outras variáveis como traços de personalidade, racionalidade, sexo e denominação religiosa. Por

fim, são comparadas as médias e as diferenças entre grupos religiões no que tange ao nível da crença na ciência.

No manuscrito 3, o processo de construção dos cenários de julgamento e decisão é descrito. O instrumento elaborado apresenta três cenários cujos objetivos envolvem decisões de natureza financeira e três cenários em que as decisões versam sobre gestão de pessoas. Além da análise de juízes, foi utilizada a técnica da análise das classes latentes como forma de se observar se os testes possuem capacidade de discriminar e classificar os indivíduos de acordo com seu padrão de resposta. Por fim, com o objetivo de identificar indícios de validade de critério, as variáveis racionalidade (com indícios de validade descritos no manuscrito 1) e crença em ciência (com indícios de validade descritos no manuscrito 2) são testadas em conjunto com traços de personalidade, sexo e experiência profissional como forma de observar a relação entre estas e as respostas tidas como corretas (racionais) presentes no teste.

A relação entre os manuscritos se consolida no manuscrito 3 e a contribuição esperada por este trabalho é que os resultados aqui encontrados sugerem a existência de indícios de validade dos instrumentos, o que pode subsidiar estudos futuros acerca de julgamento e decisão aplicado as organizações. Ao final, serão feitas considerações, além de apontar limitações e sugestão de agenda futura de pesquisas.

Referências

- Cohen, H. B. (2019). An inconvenient truth about leadership development. *Organizational Dynamics*, 48(1), 8–15. <https://doi.org/10.1016/j.orgdyn.2017.10.002>
- Farias, M., Reiman, A.K., Kahane, G., & Toledo, Z. (2013). Scientific faith: Belief in science increases in the face of stress and existential anxiety. *Journal of experimental social psychology*. 49. 1210-1213. [10.1016/j.jesp.2013.05.008](https://doi.org/10.1016/j.jesp.2013.05.008).
- Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19 (4): 25-42. DOI: [10.1257/089533005775196732](https://doi.org/10.1257/089533005775196732)

- Irwin, H. J., Drinkwater, K., & Dagnall, N. (2015). The role of doublethink and other coping processes in paranormal and related beliefs. *Journal of the Society for Psychical Research*, 79(919[2])[2], 80–97.
- Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making*, 10, 549-563. doi:10.3389/fpsyg.2013.00279
- Rutjens, B. T., Sutton, R. M., & van der Lee, R. (2018). Not all skepticism is equal: Exploring the ideological antecedents of science acceptance and rejection. *Personality and Social Psychology Bulletin*, 44(3), 384–405. <https://doi.org/10.1177/0146167217741314>
- Simon, H. A. (1955) "A behavioral model of rational choice". *The Quarterly Journal of Economics*, vol. 69, n. 1, February
- Stagnaro, M. N., Ross, R. M., Pennycook, G., & Rand, D. G. (2019). Cross-cultural support for a link between analytic thinking and disbelief in god: evidence from India and the United Kingdom. *Judgment and Decision Making*, 14(2), 179-186.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131. <http://dx.doi.org/10.1126/science.185.4157.1124>

**Manuscrito 1 - Rápido ou Devagar? – Evidências de Validade do Teste de
Reflexão Cognitiva (CRT) Para o Contexto Brasileiro**

**Fast or Slow? – Evidences of Validity of the Cognitive Reflection Test (CRT) for
the Brazilian Context.**

Thiago Augusto Costa de Olival
Universidade de Brasília

Resumo

A mensuração dos fenômenos de racionalidade possui múltiplas formas. Uma das maneiras utilizadas é com o instrumento chamado Cognitive Reflection Test (CRT) que possui 3 itens e objetiva a deliberação sobre problemas de cunho lógico matemático envolvendo a interpretação do problema e a realização de uma expressão algébrica simples. A medida sugere que quanto mais acertos, maior o traço latente (mais racional ou maior tendência a ser racional) e quanto mais erros, menor a tendência a racionalidade e em contrapartida, maior a tendência a intuição. No Brasil, este é um dos primeiros estudos objetivando avaliar as qualidades psicométricas do instrumento. Com o objetivo de avaliar os indícios de validade do instrumento para o contexto brasileiro, apresentam-se as análises e resultados decorrentes do processo de adaptação do instrumento. Em uma amostra total de 662 indivíduos foram conduzidas a análise fatorial exploratória, além da realização da TRI (Teoria de Resposta ao Item).

Palavras-chave: Cognitive Reflection Test, Racionalidade, CRT, Teoria do sistema dual

Abstract

The measurement of rationality phenomena has multiple forms. One of the ways used is with the instrument called Cognitive Reflection Test (CRT), which has 3 items and aims to deliberate on problems of mathematical logic involving the interpretation of the problem and the realization of an expression simple algebraic. The measure suggests that the more hits, the greater the latent trait (more rational or the greater tendency to be rational) and the more errors, the lower the tendency to rationality and in return, the greater the tendency to intuition. In Brazil, this is one of the first studies aiming to evaluate the psychometric qualities of the instrument. In order to assess the evidence of validity of the instrument for the Brazilian context, the analysis and results resulting from the process of adaptation of the instrument are presented. In a total sample of 662 individuals, exploratory factor analysis was conducted, in addition to the realization of IRT (Item Response Theory).

Keywords: Cognitive Reflection Test, Rationality, CRT, Dual System Theory

Introdução

Com apenas 3 itens, a proposta do Teste de Reflexão Cognitiva (CRT) é de mensurar os fenômenos racionalidade e intuitividade, ambos inseridos na teoria do sistema dual de processamento (Evans, 2008; Kahneman, 2012; Evans & Stanovich, 2013). O construto em questão é baseado na hipótese de que quanto maior a pontuação no teste, mais racional o indivíduo seria. Em contrapartida, quanto menor a pontuação, maior a probabilidade de classificação do julgamento realizado como intuitivo. Além da versão citada, existem propostas desenvolvidas por outros pesquisadores que consideram os três itens originais e incluem outros 4 itens com o mesmo contexto (Shenhav et al., 2012; Thoson & Oppenheimer, 2016). O presente estudo utilizou a versão com três itens do teste.

Com a publicação de livros como *Rápido e Devagar* (Kahneman, 2012) o Teste de Reflexão Cognitiva (CRT), desenvolvido por Frederick (2005), tem sido acessado e conhecido fora do ambiente de pesquisas acadêmicas. Embora a hipótese de conhecimento prévio das respostas possa ser uma maneira de trazer algum viés ao estudo, em uma pesquisa realizada nos Estados Unidos, Pennycook e Rand (2019) apresentam evidências de que a divulgação do instrumento não comprometeu sua validade preditiva (Bialek & Pennycook, 2017; Meyer et al., 2018; Staganaro et al., 2018; Srol, 2018).

No Brasil, ainda são necessários avanços sobre as informações referentes às qualidades psicométricas do Teste de Reflexão Cognitiva e o presente estudo tem por objetivo descrever evidências de validade do instrumento e discriminar, de maneira mais profunda, os seus atributos. Neste artigo apresentam-se a adaptação do instrumento para o Brasil, algumas observações de indícios de validade transcultural, de conteúdo e de construto do teste, além de avaliar a dificuldade e o poder de discriminação dos itens.

O Sistema Dual, Suas Possíveis Métricas e Relações

A teoria do sistema dual propõe que o processo de julgamento e tomada de decisão ocorre de duas formas, sendo uma de maneira automática/intuitiva e outra de maneira deliberada/racional. Na tentativa de medir o fenômeno, Frederick (2005) desenvolveu um instrumento composto por apenas três itens que ensejam deliberação para mensurar

racionalidade diante de questões algébricas. Dentre as respostas possíveis, há somente uma considerada como correta (racional) enquanto que outras opções diferentes do gabarito são consideradas como erradas (rápidas e intuitivas).

Os itens são de avaliação lógica, visto que objetivam a interpretação da sentença e a solução de um problema de base algébrica. Toplak et al. (2011) consideram que as questões do teste são de interesse de pesquisadores que trabalham com temas envolvendo heurísticas e preconceito, pois ao realizar o teste, uma resposta alternativa é concebida inicialmente e pode ser substituída por outra em um processo de racionalização. A suposição que envolve o erro no teste é que, ao se deparar com uma das questões, o indivíduo formule uma resposta que poderia ser equivocada e baseada em uma interpretação insuficiente do enunciado. Múltiplos estudos encontraram correlações significativas entre a alta performance no teste e: maior ceticismo em relação ao sobrenatural (Shenhav et al., 2012); alto desempenho em várias tarefas que apresentam a resposta normativa em oposição a intuição (Pennycook et al., 2015); além da correlação com medidas elevadas no teste de QI (Toplak et al., 2011) e com os traços de personalidade como neuroticismo e extroversão de maneira negativa (Cueva et al, 2016). Em pacientes com esquizofrenia, Puveendrakumaran et al. (2020) observaram que a pontuação no teste apresentou uma reflexão cognitiva intuitiva menor (baixo índice de acerto) do que em relação ao grupo controle. Os resultados indicam que os pacientes com esquizofrenia utilizaram menos o raciocínio analítico na resolução dos problemas apresentados.

O teste tem sido utilizado consideravelmente em pesquisas no mundo, mostrando seu poder preditivo em temas relacionados a tomada de decisão (Campitelli & Labollita, 2010; Hoppe & Kusterer, 2011; Neyse, 2016). Ainda há fortes evidências de que a baixa pontuação no teste seja preditora de conservadorismo quanto à atualização de probabilidades, ou seja, manter uma opinião sobre a probabilidade de um evento mesmo depois de receber novas informações (Liberali et al., 2012; Alós-Ferrer & Hügelschäfer, 2016). No campo da economia e finanças comportamentais, alguns estudos observaram relação entre a educação financeira, o teste de reflexão cognitiva e as boas escolhas de

investimentos (Neyse et al., 2016). Outra correlação do teste está associada às preferências políticas e econômicas, indicando a relação entre conservadorismo e os baixos índices no teste, enquanto que tendências libertárias estariam relacionadas aos altos índices (Pennycook, 2018).

Embora o teste proponha uma medida razão-intuição, não significa que pessoas que utilizam muito a intuição para tarefas complexas, como um enxadrista bastante intuitivo que memorizou diversas jogadas, não possam pontuar de maneira alta no teste (Pennycook et al, 2015). O teste em si poderia distinguir as pessoas que raciocinaram melhor e deram a resposta correta, mas ainda permanecem dúvidas de que as respostas erradas se tratem de erros intuitivos ou matemáticos. A suposta falta ou excesso de “intuitividade” neste domínio, conforme a hipótese de Pennycook et al. (2015), pode ocorrer em virtude da experiência com problemas matemáticos que alguns indivíduos possuem e não por que são mais ou menos intuitivos. De fato, as intuições matemáticas podem ser bem diferentes das intuições daqueles com baixa capacidade matemática. Mas, para fins de avaliação entre certo ou errado, é quase impossível distinguir, sem que exista outra medida de habilidades matemáticas como parâmetro, se o erro foi relacionado a um domínio matemático insuficiente ou se o indivíduo simplesmente não refletiu adequadamente sobre o enunciado por pura desatenção. A hipótese da relação entre as medidas de racionalidade e raciocínio matemático observada por Campitelli e Gerrans (2014) evidencia uma correlação entre o CRT e medidas de numerácia ($r=0,43$), consistente com outros estudos dessa natureza.

O que se observa é que, embora essa medida (itens do CRT) esteja relacionada ao conceito de racionalidade, tal observação poderia corresponder apenas a parte do fenômeno mais voltada ao conceito algébrico. Apesar das correlações apresentadas em alguns estudos considerando medidas que não envolvem diretamente habilidades matemáticas como a descrença no sobrenatural ou as tendências menos conservadoras, não é possível afirmar que esses mesmos indivíduos mais descrentes ou menos conservadores também não possuíssem habilidades matemáticas mais desenvolvidas, configurando talvez um viés amostral presente nos estudos citados.

Mensuração da Reflexão Cognitiva no Contexto Brasileiro

Para o contexto brasileiro, umas das versões existentes do instrumento é a mesma encontrada no livro de Daniel Kahneman, que também é utilizada em artigos científicos e trabalhos acadêmicos, mas com poucos indícios de validade. No presente estudo, realizou-se a adaptação do teste CRT ao contexto brasileiro, o que transcende somente a tradução literal dos termos ou a utilização do teste contida no livro como se tivesse sido feita por psicometristas. A presente adaptação teve como premissa identificar se os resultados refletem o contexto do teste e alguns de seus índices já observados na literatura. Também serão avaliadas algumas das qualidades psicométricas do instrumento como indícios de validade interna (convergente, discriminante, conteúdo e transcultural), a confiabilidade do instrumento e índices de dificuldade e discriminação. Os indícios de validade serão observados com o processo de tradução e adaptação, julgamento por juízes, análise fatorial exploratória e análises via TRI.

Método

Adaptação do Teste ao Contexto do Brasil

Com o intuito de evitar que os instrumentos sejam traduzidos de forma literal e que tal processo ocasione um colapso dos indícios de validade do construto, Sireci (2005) sugere que a adaptação de instrumentos entre as culturas seja realizada de maneira a considerar que os conceitos atribuídos sejam mantidos, visto que a simples tradução não garante que os termos e contextos sejam preservados.

De maneira a convergir com os pressupostos de Borsa et al. (2012), foram realizados os seguintes procedimentos no presente estudo: a (1) tradução do instrumento do idioma de origem para o idioma alvo; (2) síntese das versões traduzidas; (3) avaliação da síntese por juízes experts, (4) avaliação do instrumento pelo público-alvo; (5) tradução reversa, (6) estudo piloto e (7) avaliação da estrutura fatorial do instrumento.

Tradução do Instrumento

O processo foi realizado por um profissional com Bacharelado em Letras e Tradução, com experiência na área e que morou nos Estados Unidos por 3 anos. Adicionalmente,

outras traduções livres realizadas por mais dois indivíduos com notório saber no idioma de origem também apresentaram convergência com a tradução. Alguns dos pressupostos encontrados na literatura sobre a tradução e adaptação do instrumento, e que foram atendidos plenamente por este estudo, envolvem: a utilização de dois tradutores bilíngues (Cassepp-Borges et al., 2010); comparação com a versão traduzida do livro Rápido e Devagar encontrada no Brasil; e a consideração de que os tradutores não possuam familiaridade com o construto envolvido (Beaton et al., 2000).

Síntese do Instrumento, Avaliação Por Especialistas e Pelo Público Alvo

A versão de consenso que atendia aos itens observados por Borsa et al. (2012) apresentou algumas alterações semânticas e de tradução para melhorar a adaptação ao contexto brasileiro. O termo “*bat*” pode ser traduzido como bastão ou taco, o que não alteraria o significado do contexto da frase. A moeda utilizada no original eram US\$ dólares e foram utilizados R\$ reais, apenas para inserção do contexto nacional. O termo “*lily pads*” refere-se a uma planta chamada ninfeia, que é bem incomum em perímetros urbanos e de pouco ou nenhum uso no linguajar comum. Para o contexto, optou-se então pelo termo “ramo de lírios”, que também é a tradução possível para “*lily pads*”. A versão adotada neste estudo apresentou semelhança com as versões em português encontradas no livro publicado por Daniel Kahneman.

O grupo de juízes foi composto por 3 Bacharéis em Administração de empresas e 1 Bacharel em Economia, todos mestrandos em Psicologia do Trabalho e das Organizações, 3 Doutorandos em Psicologia e um Doutor em Psicologia, totalizando 8 indivíduos. O mesmo teste também foi aplicado em 5 alunos de graduação sem que houvesse maiores intercorrências quanto a instruções, significado e sentido das frases apresentadas. A taxa de concordância observada pelo comitê foi superior aos 90% indicados por Alexandre e Coluci (2011), considerando a proporção de itens que receberam notas 4 ou 5 em uma escala de 1 a 5, em que o 5 representa o maior grau de concordância quanto a clareza e relevância dos itens e 1 referindo-se ao menor grau.

Tradução Reversa

Muito mais do que a tradução literal já exposta anteriormente como insuficiente para garantir a qualidade integral da adaptação, Beaton et al. (2000), Sireci et al. (2005) e Borsa et al. (2012) recomendam que seja considerada a possibilidade de verificar se o contexto do instrumento original permanece, em um processo de tradução reversa. Desta forma, o instrumento foi submetido a dois universitários nativos em língua inglesa residentes nos Estados Unidos e que possuem o português como segunda língua. Os resultados entre as traduções apresentaram semelhança e o contexto de ambas as traduções não se perdeu.

Como exemplo de divergência em relação aos termos encontrada entre as traduções reversas, a expressão “taco” utilizada na questão 1 foi retraduzida como *bat* (a exemplo do original) e *pin*. Embora exista divergência no termo, o contexto se mantém sem alteração significativa da frase e sem distanciamento do conceito (pino, taco ou bastão). No item 3, o termo adaptado de *lily pads* foi retraduzido como ramo de lírios ou simplesmente lírios, também sem alterar o contexto do item.

A versão final do instrumento permaneceu conforme abaixo:

1 – Um taco e uma bola custam RS 1,10 no total. O taco custa um real a mais que a bola. Quanto custa a bola?

2 – Se 5 máquinas demoram 5 minutos para fazer 5 ferramentas, quanto tempo leva para 100 máquinas fabricarem 100 ferramentas?

3 – Em um lago, há um ramo de lírios. Todos os dias, o ramo dobra de tamanho. Se leva 48 dias para o ramo cobrir o lago todo, quanto tempo levará para que o ramo cubra metade do lago?

Participantes

O instrumento foi aplicado entre o final do ano de 2019 e início do ano de 2020, via *google forms* em comunidades de uma rede social, além de estar disponível na página pessoal do pesquisador. O processo de aplicação e coleta de dados totalizou 662 observações. Tendo em vista que a pesquisa possuía *link* estático, não é possível saber

quantas pessoas de fato acessaram sem concluir o questionário. Os dados da amostra estão apresentados na Tabela 1.

Tabela 1

Caracterização dos participantes da pesquisa

Características	Distribuição	
	N	%
Sexo	Masculino	293 44%
	Feminino	369 56%
Escolaridade	Nível Médio	75 11%
	Superior-Graduação	207 31%
	Superior -Pós-Graduação -Especialização	332 50%
	Superior – Mestrado	39 6%
	Superior - Doutorado	9 1%
Idade	18-25	40 6%
	26-30	57 9%
	31-35	121 18%
	36-40	175 26%
	41-45	107 16%
	46-50	68 10%
	51-55	45 7%
	56-60	27 4%
	61-65	14 2%
	66-70	6 1%
71-75	2 0%	
Região	Centro Oeste	162 24%
	Nordeste	34 5%
	Norte	13 2%
	Sudeste	399 60%
	Sul	54 8%

Análise dos Dados

Os dados avaliados da amostra não apresentaram normalidade, além da inexistência de casos extremos univariados ou dados omissos. Também não foram observados casos extremos multivariados ao se analisar a distância de *Mahalanobis*. Embora exista a preocupação no que tange aos pressupostos, as recomendações de Hair et al, (2009) são de que o tamanho da amostra ($n > 200$) e a inexistência de heterocedasticidade possam ser indicativos de que a não normalidade univariada dos dados não acarrete maiores problemas para o prosseguimento das análises.

Resultados

Os resultados são apresentados considerando as análises realizadas para aferir os indícios de validade do instrumento. Os indícios de validade de construto serão analisados por meio dos resultados das análises fatoriais exploratória e confirmatória, bem como os indícios de validade convergente e discriminante (por meio da confiabilidade composta e da

variância média extraída), ambas calculadas a partir dos resultados da análise fatorial exploratória. Outro indício de validade de construto observado diz respeito aos resultados da TRI, que também apontam evidências sobre a validade da estrutura interna do instrumento. Por fim, foram observadas as qualidades de discriminação e dificuldade dos itens e do teste como um todo, mediante suas curvas de informação e característica.

Análise Fatorial

De maneira a avaliar a consistência interna do instrumento, Borsa et al. (2012), sugerem a realização de análises fatorial exploratória e análise fatorial confirmatória como uma das formas de avaliar indícios de validade de construto. A amostra utilizada para ambas as análises atende plenamente ao pressuposto da quantidade de casos por item (Hair et al., 2009; Pasquali, 2010; Damásio, 2012; Tabachnick & Fidel, 2014) que varia na literatura entre 5 a 15 observações.

Análise Fatorial Exploratória

A extração de fatores considerando itens dicotômicos já teve observações que impactavam nos resultados, mas hoje existem soluções analíticas que permitem submeter tais escalas a análise, reduzindo os problemas de resultados enviesados (Laros, 2012). Uma análise fatorial exploratória em uma amostra com 662 indivíduos foi conduzida pelo executável *Factor* versão 10.10.03 e com o intuito de avaliar a fatorabilidade do instrumento, foi realizado o teste de Kaiser-Meyer-Olkin - KMO e o teste de esfericidade de Barlett. Para fins de adequação do instrumento a realização da análise fatorial exploratória, são esperados que além da significância no teste de esfericidade, o KMO seja superior a 0,5 (Hair et al, 2009) como patamar mínimo. Outros autores como Hutcheson e Sofroniou (1999) indicam valores acima de 0,70 como parâmetro de bom índice para o KMO. Para esta análise foi observado um KMO próximo ao limite para que seja considerado como de razoável a bom ($KMO=0,72$) além da significância no teste de esfericidade ($X^2=868,1$, $p>0,00$). Esses fatos indicam que poderia haver alguma adequação do modelo a análise fatorial.

O modelo apresentou uma estrutura platicúrtica nos 3 itens ($q_1 = -1,76$; $q_2 = -1,99$; $q_3 = -1,97$) e assimetria levemente inclinada à direita ($q_1 = 0,50$; $q_2 = -0,08$; $q_3 = 0,16$). Os fatores foram extraídos utilizando os métodos *Robust Diagonally Weighted Least Squares* e *Minimum Rank Factor Analysis* via análise paralela (Damásio, 2012; Machado et al., 2014; Hernandez, 2015; Gomes et al., 2016) e os resultados obtidos foram semelhantes entre os métodos. A indicação quanto a rotação de fatores nas ciências humanas e sociais é para a utilização de rotações oblíquas, visto a impossibilidade de isolar os fatores integralmente e o pressuposto de relação entre eles (Tabachnick & Fidel, 2014). Em todos os modelos, a estrutura unifatorial foi corroborada.

A partir da análise preliminar, a proporção da variância explicada pelo instrumento foi de 0,77, com apenas uma dimensão e autovalor acima de 1. As cargas fatoriais observadas ($q_1 = 0,78$; $q_2 = 0,76$; $q_3 = 0,89$) foram elevadas, assim como as comunalidades, que foram superiores a 0,6 e inferiores que 0,9. Sobre a confiabilidade da escala, observou-se um Alfa de Cronbach de 0,71, valor este indicado pela literatura (Hair et al., 2009; Maroco & Garcia-Marques, 2013; Alexandre et al., 2013; Souza et al., 2017) como mínimo aceitável. O instrumento apresentou estrutura unifatorial com cargas fatoriais altas, índices de adequação da amostra e fatorabilidade baixos, mas aceitáveis, de acordo com a literatura. Os demais índices de ajuste do modelo são apresentados conforme dados transcritos na Tabela 2.

Tabela 2

Relação dos índices da análise

	<i>Índices</i>	
	<i>Observados</i>	<i>Esperado</i>
χ^2	768,23	-
<i>Gl</i>	3	-
<i>P-valor</i>	0,000	<0,05
χ^2/GL	289,37	<5
<i>CFI</i>	1,00	>0,9
<i>BIC</i>	38,97	-
<i>RMSEA</i>	0,00	<0,08
<i>GFI</i>	1,00	>0,9
<i>CC</i>	0,85	0,70
<i>VME</i>	0,66	0,50

Nota: índices esperados conforme Hair et al (2009)

O ajuste do modelo também demonstrou indicies satisfatórios ($RMSEA = 0$; CFI e $TLI = 1$), o que estaria de acordo com os parâmetros sugeridos, porém podem indicar um

superajuste devido a influência da quantidade reduzida de itens do instrumento (Hu & Bentler, 1999; Kenny & McCoach, 2003). A correlação média ao quadrado entre os itens foi de 0,43, sendo menor do que a variância média extraída ($VME = 0,66$), indicando uma evidência de validade discriminante.

TRI (Teoria de Resposta ao Item)

Com o propósito de aprimorar os indícios de validade do instrumento, também foi realizada a TRI do instrumento, cuja função é avaliar a discriminância e dificuldade dos itens. Paquali & Primi (2003) apontam algumas das vantagens da TRI como a possibilidade de prever o desempenho do sujeito em uma tarefa, de acordo com as aptidões ou traços latentes (identificados na TRI como a letra grega *tetha*), além da descrição matemática monotômica do desempenho, chamada de curva característica do item. Tendo em vista a necessidade de avaliar o *tetha* dos sujeitos, Couto & Primi (2011) sugerem que as análises da TRI poderiam ajudar os profissionais a aprimorar a capacidade de realizar julgamentos sobre a qualidade dos instrumentos com base em princípios calcados no método científico. Considerando as observações de Pasquali (2009), a psicometria auxiliaria a criação de instrumentos psicológicos, além de viabilizar ao pesquisador a verificação se o teste mede o que se propõe a medir.

Conforme sintetizam Couto & Primi (2011), o índice de discriminação do item representado pela letra “a”, apresenta seu valor pela inclinação da Curva Característica dos Itens em relação ao eixo das abscissas, o parâmetro “b” é denominado índice de dificuldade do item e seu valor é dado pelo ponto perpendicular ao eixo das abscissas, representando o valor de (θ) quando a probabilidade de o sujeito acertar ao item é de 50%, enquanto o “c” é a probabilidade do acerto ao acaso (chute).

O instrumento foi disponibilizado para que os indivíduos respondessem de maneira aberta, sem a presença de itens de múltipla escolha, sendo que as respostas deveriam ser transcritas de forma numérica ou textual. Assim, o acerto ao acaso (c) foi desprezado por não fazer muito sentido neste momento, e o modelo de 2PL foi assumido para fins de conhecimento da dificuldade e discriminação dos itens. Para realizar a análise, as respostas

foram transformadas em dados binários, ocasionando em 8 padrões de resposta possíveis, considerando uma matriz composta da quantidade total de itens corretos e incorretos e quais itens compõem o padrão (apenas o item 1 correto e demais incorretos, item 1 e 3 corretos e item 2 incorreto, etc.).

Quando avaliada a correlação entre os itens considerando a técnica da TRI, observa-se que individualmente os itens tem grande índice de correlação bisserial variando de 0,78 a 0,82, em relação ao resultado total do teste. A correlação bisserial estaria vinculada a relação do item com o conjunto dos itens, sendo sugerido que os valores observados sejam superiores a 0,3 (Pasquali, 2018). O Alfa de Cronbach apresentou índice semelhante ao observado na análise fatorial exploratória ($\alpha=0,71$) e um pouco inferior ao observado na análise fatorial confirmatória ($\alpha =0,79$). A hipótese de exclusão dos itens ocasionaria uma sensível redução no indicador, conforme se observa na tabela 3. Ademais, o instrumento apresentou um índice de informação de 7,51 (99,73% considerando -3, 3) para um $AIC = 2382,54$, $BIC = 2409,51$ e $-2 \text{ Log Likelihood} = -1185,27$, considerando o modelo com três itens como o mais adequado, se comparado com um modelo com apenas um ou dois itens.

Tabela 3

Índices dos itens do modelo

	Alfa de Cronbach	Discriminação (a)	Dificuldade (b)	Correlação Bisserial	Carga fatorial TRI	Acerto	Erro	M	SD	X^2
Q1	0,64	2,110	0,393	0,78	0,78	38%	62%	0,38	0,49	358,89
Q2	0,65	1,965	-0,065	0,78	0,76	52%	48%	0,52	0,50	385,38
Q3	0,56	3,456	0,111	0,82	0,90	46%	54%	0,46	0,50	176,01

Nota: Medidas do α considerando a exclusão do item e X^2 considerando $p>0,001$. O α da escala foi de 0,71.

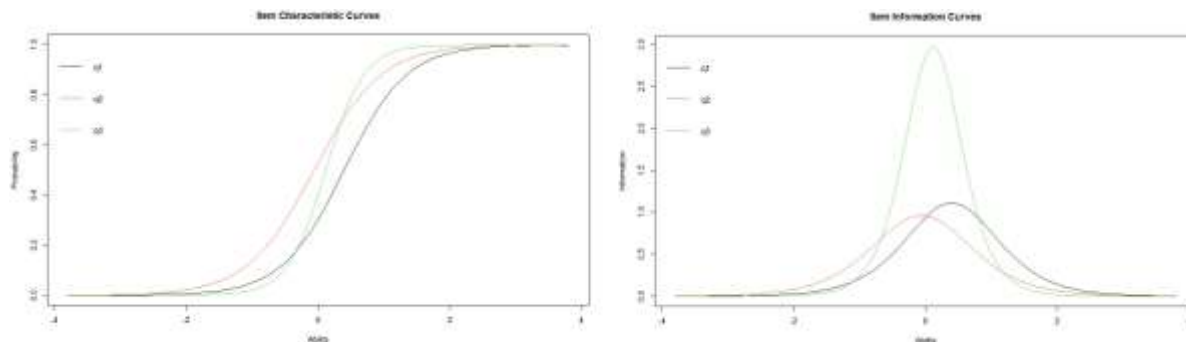
Comparando com o modelo de 1PL, observa-se que o AIC (2383,89 para 1PL) e $-2 \text{ Log Likelihood}$ (-1189,43 para 1PL) são maiores e o BIC (2404,84 para 1PL) é menor, mas os três índices encontram-se muito próximos ao resultado observado no modelo de 2PL. Por meio da análise do qui quadrado com base no $-2 \text{ Log Likelihood}$, observou-se que as diferenças são significantes, e que o modelo 2PL apresenta melhoria na análise dos dados.

Ao se observar a curva de informação (representação da discriminação) e a curva de característica (representação da dificuldade), percebe-se que os itens em questão têm faixas de discriminação e dificuldade parecidas, se comportando de maneira bastante similar

no que tange a sua qualidade e faixa de mensuração. A exceção ao padrão é a discriminação do item 3, conforme se observa na figura 1.

Figura 1

Comparativo entre padrão CCI (Curva Característica do Item) e CII (Curva de Informação do Item)



Comparação Entre Resultados Brasileiros e Estudos de Outros Países

Os resultados da aplicação do instrumento no Brasil apresentaram consideráveis semelhanças com resultados encontrados em estudos internacionais no que tange a estrutura fatorial, sua carga e outros indicadores. As médias e o desvio padrão por item convergem com as encontradas pelo estudo de Campitelli e Gerrans (2014) e Kirkegaard e Nordbjerb (2015).

Frederik (2005) conduziu seu estudo inicial com universitários e observou a distribuição das pontuações no teste, considerando a somatória dos acertos das questões entre zero (baixo) e três (alto). Na tabela abaixo são comparados os resultados de Frederick (2005), Kirkegaard e Nordbjerb (2015) e os observados no presente estudo no que tange as médias e padrão de escore.

Tabela 4

Comparativo Entre Estudos

Local da coleta	M	n	%			
			0	1	2	3
MIT	2,18	61	7%	16%	30%	48%
Princeton University	1,63	121	18%	27%	28%	26%
Boston Fireworks Display	1,53	195	24%	24%	26%	26%
Carnegie Mellon Univ	1,51	746	25%	25%	25%	25%
Harvard University	1,43	51	20%	37%	24%	20%
Univ. of Michigan: Ann Arbor	1,18	1267	31%	33%	23%	14%
Estudos na internet	1,10	525	39%	25%	22%	13%
Bowling Green University	0,87	52	50%	25%	13%	12%
Univ. of Michigan: Dearborn	0,83	154	51%	22%	21%	6%
Michigan State Univ.	0,79	118	49%	29%	16%	6%
Universisty of Toledo	0,57	138	64%	21%	10%	5%
Total (Frederick (2005)	1,24	3428	33%	28%	23%	17%
Kirkegaard & Nordbjerb (2015)	1,11	72	-	-	-	-
Amostra deste estudo	1,36	662	34%	21%	21%	24%

Nota. Considerando o objetivo do presente estudo, é importante ressaltar que o instrumento apresentou resultados similares aqueles encontrados em outros países (ver Frederick, 2005; Campitelli & Gerrans, 2014; Kirkgaard & Nordbjerd, 2015).

A comparação com um modelo de 2PL via TRI aplicado por Primi et al., (2016) também apresentou resultados similares ao deste estudo, em que se observou que os três itens possuem discriminação alta ($q1_a=1,7$; $q2_a=1,48$; $q3_a=3,05$) e dificuldade de fácil a média ($q1_b=0,26$; $q2_b=0,70$; $q3_b=0,08$) para um índice de informação de 6,10.

Dentre os achados deste estudo, destacam-se: o índice de concordância na avaliação dos juízes que foi superior ao referenciado na literatura; a fatorabilidade do instrumento que corrobora a hipótese de que os itens se agruparam de maneira a possibilitar a mensuração do fenômeno; a confiabilidade da escala que, embora limítrofe ao sugerido, se alinha aos demais estudos já realizados; a confiabilidade composta e a variância média extraída dentro dos parâmetros esperados - ambos índices atrelados aos indícios de validade convergente e discriminante; além dos indicadores de dificuldade e discriminação dos itens, bem como as curvas característica e de informação do instrumento extraídos via TRI que indicaram medidas de traço latente dos sujeitos. Os resultados observados apontam múltiplos indícios de validade interna, que serão discutidos de acordo com os pressupostos específicos de cada evidência.

Discussão

Ao que se demonstra, existem diversas similaridades dos resultados observados neste estudo quando comparados com os achados de estudos internacionais (ver Campitelli & Gerrans, 2014; Kirkegarrd & Nordbjerb, 2015). O processo de adaptação ao contexto brasileiro foi realizado em conformidade com os pressupostos elencados (Beaton et al., 2000; Sireci et al., 2005; Borsa et al., 2012; Cassep-Borges et al., 2010). Desta forma, sugere-se que a adaptação e tradução permitiram avaliar indícios de validade transcultural com muita pertinência. Tais resultados apresentam evidências de que o processo da adaptação transcultural não alterou a essência do instrumento. Considera-se que existem indícios para subsidiar a utilização do instrumento no contexto brasileiro.

Existem indícios de validade de conteúdo observados em consonância com as

sugestões de Alexandre e Coluci (2011), visto que o grau de convergência entre as análises de 8 juízes e especialistas foi elevada e o $IVC = 100\%$, considerando uma escala de 1 a 5. O cálculo proposto é representado pela somatória de respostas com escore 4 ou 5 e dividido pelo número total de respostas. Embora este cálculo seja costumeiramente apresentado com escalas de 1 a 4, Alexandre e Coluci (2011) não sugerem óbices quanto a utilização da escala com 5 pontos. Os parâmetros avaliados foram clareza, pertinência e relevância dos itens.

A confiabilidade do instrumento ($\alpha = 0,71$; $\lambda^2 = 0,72$) também mantém média semelhante aos estudos internacionais, embora ainda seja motivo de atenção pois é considerada como próxima ao limite considerável como aceitável para os pressupostos gerais de confiabilidade, observada em múltiplos artigos sobre o tema (Maroco & Garcia-Marques, 2013; Souza et al, 2017). A confiabilidade composta está adequada as indicações de Hair et al., (2009) e Valentini e Damasio (2016), pois apresentam valor de 0,85. Ao mesmo tempo, a VME extraída entre os itens e o fator também se encontra superior ao valor do pressuposto, o que sugere evidências de validade convergente. Ainda, observam-se evidências de validade discriminante visto que a VME é maior do que as correlações dos itens ao quadrado. Observa-se que o modelo unifatorial está bem ajustado e com cargas fatoriais robustas para cada um dos itens, o que implicaria em aceitar que apesar de suas limitações em alguns índices, o modelo proposto é adequado, mas com ressalvas. Uma hipótese para a inadequação do modelo no que tange a relação entre os graus de liberdade e X^2 pode ser devido a quantidade de itens presentes no instrumento (que são apenas 3). A mesma hipótese é discutida por Valentini e Damasio (2016) sobre os possíveis impactos da quantidade de itens nos coeficientes alfa de Cronbach, VME e confiabilidade composta.

Outro indício de validade sugerido por Pasquali (2007) diz respeito a observação de validade de construto via TRI, visto que com análise é possível verificar o traço latente via função de informação. As orientações prescritas por Pasquali (2018) quanto a interpretação da TRI, no que diz respeito ao “a” (discriminação), ao “b” (dificuldade) e no que tange a avaliação da curva característica do item, são para considerar a probabilidade de 0,5 para

respostas corretas ao acaso. Os itens com “a” acima de 1,70 possuem alta discriminação, isto é, alta capacidade de discriminar sujeitos em relação aos níveis do traço latente avaliado (Andrade et al. 2000) e o que se observou é que todos os itens do instrumento possuem índices satisfatórios sobre este quesito. O item 3 (q3) possui a melhor discriminação, o que demonstra sensibilidade considerável para a mensuração de mesmo nível de *tetha* em múltiplos indivíduos (Sartes & Souza-Formigoni, 2013).

No que tange a dificuldade, ao serem consideradas a expectativa de valores esperados (-2 a +2), conforme Baker (2001), os itens da escala poderiam ser considerados como “medianos”, sendo o item 1 (q1) o mais próximo do parâmetro “difícil” e o item 2 (q2) como o mais próximo de “fácil”, tendo como parâmetro somente a comparação dentro da escala e para esta amostra. Importante ressaltar que assim como no estudo de Primi et al., (2016), a escala apresenta índices satisfatórios para a diferenciação de *tetha* de indivíduos que não estejam entre as extremidades mais baixas ou mais altas do traço latente.

Os índices de dificuldade e discriminação demonstram que os itens possuem boa qualidade para a medição do traço latente de maneira mais central na escala, além de apresentar confiabilidade e estrutura fatorial. Como é uma escala com quantidade reduzida de itens, pode ser facilmente incluída em pesquisas sem que se eleve substancialmente o tempo ou a quantidade de perguntas de um questionário. O instrumento é constantemente citado na literatura internacional e este estudo apresenta alguns dos primeiros indícios de validade para o contexto brasileiro, o que demonstra que estudos adicionais ainda precisam ser feitos para que sejam observadas correlações de mesma natureza das elencadas em outros países.

É possível concluir que os indícios de validade observados apontam que o construto racionalidade pode ser medido pelo instrumento, mas os resultados sugerem que algumas das ressalvas ou inadequações a pressupostos estatísticos estariam relacionados a quantidade reduzida de itens da escala. Ao mesmo tempo, permanecem inconclusivas as correlações com a numerácia (Campitelli & Gerrans, 2014) ou outros construtos que tenham como objetivo medidas de racionalidade, pois estas não foram alvo deste estudo. O traço

latente observado no teste é sugerido por Frederick (2005) como uma medida da resistência a responder as questões de maneira impulsiva e até o momento não existem evidências que não seja somente uma medida de habilidade aritmética, apesar da relação entre os conceitos.

Como limitação do estudo, destaca-se a possibilidade de existir um viés no que tange ao conhecimento prévio ou a possibilidade de consulta das respostas na internet. Não foi possível saber se os indivíduos realizaram algum tipo de consulta ou se já conheciam as respostas. Para atenuar essa constatação, os resultados e médias observadas foram semelhantes às aquelas observadas em outros estudos. Os possíveis impactos do conhecimento prévio dos itens, caso existam, devem ter se apresentado com o mesmo padrão dos artigos citados neste trabalho.

Em estudos futuros, podem ser feitas propostas de incremento do instrumento utilizando também testes de raciocínio lógico de outras categorias como sequências lógicas ou deduções/ induções, visto que na atual proposta ele contém somente itens de cunho lógico matemático. Uma das hipóteses de agenda futura é que a quantidade de itens poderia influenciar nos indicadores da escala, fato este que tem se mostrado uma tendência em alguns artigos com alguma medida de razão e intuição, visto que os instrumentos têm sido utilizados com a escala original suplementada de itens adicionais de natureza similar. Sugere-se comparar os dados de validade entre versões com três e mais itens.

Referências

- Alexandre, N. M. C., & Colucci, M. Z. O. (2011). Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Ciência & Saúde Coletiva*, 16(7), 3061-3068. <https://doi.org/10.1590/S1413-81232011000800006>
- Alexandre, N. M. C., Gallasch, C. H., Lima, M. H. M., & Rodrigues, R. C. M. (2013). A confiabilidade no desenvolvimento e avaliação de instrumentos de medida na área da saúde. *Revista Eletrônica De Enfermagem*, 15(3), 800-7. <https://doi.org/10.5216/ree.v15i3.20776>

- Alós-Ferrer, C., & Hügelschäfer, S. (2015). Faith in Intuition and Cognitive Reflection. *Journal of Behavioral Economics*. 64. 10.1016/j.socec.2015.10.006.
- Amorim, L. (2015). Análise fatorial confirmatória e modelos com equações estruturais: Um Tutorial usando Software Estatístico. 10.13140/RG.2.1.1520.4960.
- Andrade, D. F., Valle, R. C., & Tavares, H. R. (2000) *Introdução à teoria da resposta ao ítem: conceitos e aplicações*. Sinapse.
- Baker, F. B. (2001) *The Basics of Item Response Theory*. 2^a. ed. United States of America: ERIC Clearinghouse on Assessment and Evaluation,
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186-3191.
- Białek, M., & Pennycook, G. (2018). The Cognitive Reflection Test is robust to multiple exposures. *Behavior Research Methods*. 50. 1953–1959. 10.3758/s13428-017-0963-x.
- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Adaptação e Validação de Instrumentos Psicológicos entre Culturas: Algumas Considerações. *Paidéia*, 22(53), 423–432. <https://doi.org/10.1590/S0103-863X2012000300014>
- Brañas-Garza, P., Kujal, P., & Lenkei, B. (2015). 15-25 Cognitive Reflection Test: Whom, How, When. *Journal of Behavioral and Experimental Economics*. 82. 10.1016/j.socec.2019.101455.
- Brown, T. A. (2006). *Confirmatory factor analysis for Applied Research*, 1st Edition, New York, The Guilford Press
- Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Mem Cognit*. 2014 Apr;42(3):434-47. doi: 10.3758/s13421-013-0367-9. PMID: 24132723.
- Campitelli, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making*. 5. 182-191.

- Cassepp-Borges, V., Balbinotti, M. A. A., & Teodoro, M. L. M. (2010). Tradução e validação de conteúdo: Uma proposta para a adaptação de instrumentos. In L. Pasquali, *Instrumentação psicológica: Fundamentos e práticas*. Porto Alegre: Artmed.
- Christensen, A. P., & Golino, H. (2021). On the equivalency of factor and network loadings. *Behavior research methods*. <https://doi.org/10.3758/s13428-020-01500-6>
- Comrey A., L., & Lee H., B. (2013). *A first course in factor analysis (2nd ed.)*. New York, NY: Psychology Press.
- Couto, G., & Primi, R. (2011). Teoria de resposta ao item (TRI): Conceitos elementares dos modelos para itens dicotômicos. *Boletim de Psicologia*, 61(134), 1-15.
- Cueva, C., Iturbe-Ormaetxe, I., Mata-Pérez, E., Ponti, G., Sartarelli, M., Yu, H., & Zhukova, V. (2016). Cognitive (ir)reflection: New experimental evidence. *Journal of Behavioral and Experimental Economics*, 64, 81–93. <https://doi.org/10.1016/j.socec.2015.09.002>
- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, 11(2), 213-228.
- Evans J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual review of psychology*, 59, 255–278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223–241. <https://doi.org/10.1177/1745691612460685>
- Field, A. P. (2009). *Discovering statistics using SPSS*. Thousand Oaks, CA: Sage.
- Frederick, S. (2005). "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives*, 19 (4): 25-42. DOI: 10.1257/089533005775196732
- Gomes, C. A. O., Soares, A. B., Mourão, L., & Hernandez, J. A. E. (2016). Inventário de Estilos de Enamoramento: construção e validação. *Avaliação Psicológica*, 15(2), 151-159. Recuperado em 03 de fevereiro de 2021, de
- Hair, Jr., J.F., Willian, B., Babin, B. & Anderson, R.E. *Análise multivariada de dados*. 6.ed. Porto Alegre: Bookman, 2009.

- Hambleton, R. K., Merenda, P.F., & Spielberger, C.D., (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum.
- Hoppe, E. I., Kusterer D.(2011), J. Behavioral biases and cognitive reflection, *Economics Letters*, Volume 110, Issue 2, 2011, Pages 97-100, ISSN 0165-1765, <https://doi.org/10.1016/j.econlet.2010.11.015>.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55
- Hutcheson, G. D. & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. London: Sage Publications
- Kahneman, D. (2012). *Rápido e devagar, duas formas de pensar*. Rio de Janeiro: Editora Objetiva.
- Kenny, D., A., & McCoach, D., B., (2003) *Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling*, *Structural Equation Modeling: A Multidisciplinary Journal*, 10:3, 333-351, DOI: 10.1207/S15328007SEM1003_1
- Kirkegaard, E. O. W., & Nordbjerg, O. (2015). Validating a Danish translation of the International Cognitive Ability Resource sample test and Cognitive Reflection Test in a student sample. *Open Differential Psychology*. 10.26775/ODP.2015.07.31.
- Laros, J. (2012). O Uso da Análise Fatorial: Algumas Diretrizes para Pesquisadores. Em: L.Pasquali (org), *Análise fatorial para pesquisadores*. Petrópolis: Vozes
- Liberali, J., Reyna, V., Furlan, S., Stein, L., & Pardo, S. (2012). Individual Differences in Numeracy and Cognitive Reflection, with Implications for Biases and Fallacies in Probability Judgment. *Journal of behavioral decision making*. 25. 361-381. 10.1002/bdm.752.
- Machado, W. L., Hauck, N., Teixeira, M. A. P., & Bandeira, D. R. (2014). Análise de teoria de resposta ao item de marcadores reduzidos da personalidade. *Psico*, 45(4), 551-558.
- Maroco, J., & Garcia-Marques, T. (2013). Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia*. 4. 10.14417/lp.763.

- Mata, A., Fiedler, K.s., Ferreira, M., & Almeida, Ti. (2013). Reasoning about others' reasoning. *Journal of Experimental Social Psychology*. 49. 486–491.
10.1016/j.jesp.2013.01.010.
- Meyer, A., & Zhou, E., & Frederick, S. (2018). The non-effects of repeated exposure to the cognitive reflection test. *Judgment and Decision Making*. 13. 246-259.
- Neyse, L., Bosworth, S., Ring, P. *et al.*(2016). Overconfidence, Incentives and Digit Ratio. *Sci Rep* 6, 23294). <https://doi.org/10.1038/srep23294>
- Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, 72(1), 147–152
- Pasquali, L. (2009) *Psicometria: Teoria dos testes na psicologia e na educação*. 3. ed. Petrópolis, RJ: Vozes,
- Pasquali, L. (2010). *Instrumentação Psicológica: fundamentos e práticas*. Porto Alegre: Artmed
- Pasquali, L. (2018) *Teoria de Resposta ao Item – Teoria, procedimentos e aplicações*. 3 ed. Curitiba. Appris
- Pasquali, L., & Primi, R. (2003). Fundamentos da teoria da resposta ao item: TRI. *Avaliação Psicológica*, 2(2), 99-110.
- Pennycook, G. (2018). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*. 188.
10.1016/j.cognition.2018.06.011.
- Pennycook, G., & Rand, D. G. (2019). Cognitive Reflection and the 2016 U.S. Presidential Election. *Personality and Social Psychology Bulletin*, 45(2), 224–239. <https://doi.org/10.1177/0146167218783192>
- Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making*, 10, 549-563. doi:10.3389/fpsyg.2013.00279

- Pennycook, G., Cheyne, J.A., Koehler, D.J. *et al.*(2016) Is the cognitive reflection test a measure of both reflection and intuition?. *Behav Res* 48, 341–348 (2016).
<https://doi.org/10.3758/s13428-015-0576>
- Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2016). Atheists and agnostics are more reflective than religious believers: Four empirical studies and a meta analysis. *PLoS ONE*, 11(4). <https://doi.org/10.1371/journal.pone.0153039>
- Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016) The Development and Testing of a New Version of the Cognitive Reflection Test Applying Item Response Theory (IRT). *J. Behav. Dec. Making*, 29: 453– 469. doi: 10.1002/bdm.1883.
- Primi, R., & Almeida, L.S. (1998). Considerações sobre a análise factorial de itens com resposta dicotômica. *Psicologia: Teoria, Investigação e Prática*, 3, 225-234.
- Puveendrakumaran, P., Fervaha G, Caravaggio, F., & Remington, G. Assessing analytic and intuitive reasoning using the cognitive reflection test in young patients with schizophrenia. *Psychiatry Res.* 2020. doi: 10.1016/j.psychres.2019.112683.
- Saidi, S., Purwanto, P., Asari, A & Sudirman, S. (2019). Study of student error thinking in solving Cognitive Reflection Test (CRT) required from Dual Process Theory (DPT). *Journal of Physics: Conference Series*. 1157. 032088. 10.1088/1742-6596/1157/3/032088.
- Sartes, L. M. A., & Souza-Formigoni, M. L. O. (2013). Avanços na psicometria: da Teoria Clássica dos Testes à Teoria de Resposta ao Item. *Psicologia: Reflexão e Crítica*, 26(2), 241-250. <https://doi.org/10.1590/S0102-79722013000200004>
- Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, 141(3), 423–428. <https://doi.org/10.1037/a0025391>
- Simon, H. A. (1955) "A behavioral model of rational choice", *The Quarterly Journal of Economics*, vol. 69, n. 1, February: 99-118, compiled in, and quoted from, Simon (1957: 241-260).

- Sireci, S. G. (2005). Using bilinguals to evaluate the comparability of difference language versions of a test. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 117-138). Mahwah, NJ: Lawrence Erlbaum.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22. doi:10.1037/0033-2909.119.1.3
- Souza, A. C., Alexandre, N. M. C., & Guirardello, E. B. (2017). Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. *Epidemiologia e Serviços de Saúde*, 26(3), 649-659. <https://dx.doi.org/10.5123/s1679-49742017000300022>
- Šrol, J.(2018). These Problems Sound Familiar to Me: Previous Exposure, Cognitive Reflection Test, and the Moderating Role of Analytic Thinking. *Studia psychologica*. 60. 195-208. 10.21909/sp.2018.03.762.
- Stagnaro, M. N., Pennycook, G., & Rand, D. G. (2018). "Performance on the Cognitive Reflection Test is stable across time," *Judgment and Decision Making, Society for Judgment and Decision Making*, vol. 13(3), pages 260-267, May.
- Stagnaro, M., & Pennycook, G. (2018). Cognitive Reflection is a Stable Trait. *SSRN Electronic Journal*. 10.2139/ssrn.3115809.
- Tabachnick, B. G., and Fidell, L. S. (2014). *Using Multivariate Statistics*, 6th ed. Harlow, UK: Pearson.
- Thomson, K. S., & Oppenheimer, D., M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, 11(1), 99–113.
- Toplak, M., West, R., & Stanovich, K. (2013). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*. 20. 147-168. 10.1080/13546783.2013.844729.
- Toplak, M.E., West R.F., & Stanovich K.E.,(2011) The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Mem Cognit*.Oct;39(7):1275-89. doi: 10.3758/s13421-011-0104-1. PMID: 21541821.

- Travers, E., Rolison, J. J., & Feeney, A, (2016) The time course of conflict on the Cognitive Reflection Test, *Cognition*, 150, 2016, Pages 109-118, ISSN 0010-0277, <https://doi.org/10.1016/j.cognition.2016.01.015>.
- Valentini, F, & Damásio, B. F. (2016). Variância Média Extraída e Confiabilidade Composta: Indicadores de Precisão. *Psicologia: Teoria e Pesquisa*, 32(2), e322225. Epub October 27, 2016. <https://doi.org/10.1590/0102-3772e322225>
- Yilmaz, O., & Saribay, S. A. (2018). Lower Levels of Resistance to Change (but not Opposition to Equality) Is Related to Analytic Cognitive Style. *Social Psychology*, 49, 65–7

Apêndice A – Teste de Reflexão Cognitiva (*Cognitive Reflection Test – CRT*)

1 – Um taco e uma bola custam RS 1,10 no total. O taco custa um real a mais que a bola.

Quanto custa a bola?

2 – Se 5 máquinas demoram 5 minutos para fazer 5 ferramentas, quanto tempo leva para 100 máquinas fabricarem 100 ferramentas?

3 – Em um lago, há um ramo de lírios. Todos os dias, o ramo dobra de tamanho. Se leva 48 dias para o ramo cobrir o lago todo, quanto tempo levará pra que o ramo cubra metade do lado?

**Manuscrito 2 - Uma Vela na Escuridão: Evidências de Validade da Escala de
Crença em Ciência (BISS) Para o Contexto Brasileiro.**

**A Candle in the Dark: Evidence of Validity of the Belief in Science's Scale
(BISS) for the Brazilian Context.**

Thiago Augusto Costa de Olival

Universidade de Brasília

Resumo

A escala de crença em ciência (Believe in Science Scale –BISS) tem como objetivo mensurar o fenômeno desta crença de acordo com a concordância acerca de 10 afirmações relacionadas a ciência e a concordância com sua visão de mundo. Ao contrário da crença, a descrença científica apresenta relacionamentos com algumas crenças religiosas, embora a princípio não existam óbices quanto a este suposto pensamento duplo de crer em ciência e crer no sobrenatural. Desta forma, este manuscrito busca avaliar as qualidades psicométricas do instrumento para o contexto brasileiro, focado principalmente nos indícios de validade. Apresenta-se uma versão traduzida do instrumento que contou com a adaptação ao contexto brasileiro, além da realização da análise fatorial exploratória (n = 662), correlação entre crença e traços de personalidade, racionalidade, religião, sexo, idade e escolaridade, além de realizar a comparação entre as médias da crença em ciência de grupos religiosos.

Palavras-chave: Crença em ciência, Ciência e religião, Cientificismo

Abstract

The belief scale in science (Believe in Science Scale - BISS) aims to measure the phenomenon of this belief according to the agreement about 10 statements related to science and the agreement with their worldview. Contrary to belief, scientific disbelief has relationships with some religious beliefs, although at first there are no obstacles to this supposed dual thinking of believing in science and believing in the supernatural. In this way, this manuscript seeks to evaluate the psychometric qualities of the instrument for the Brazilian context, focused mainly on the signs of validity. A translated version of the instrument is presented, which included adaptation to the Brazilian context, in addition to carrying out exploratory factor analysis (n = 662), correlation between belief and personality traits, rationality, religion, sex, age and education, in addition to to carry out the comparison between the averages of the belief in science of religious groups.

Keywords: Science Belief, Science and Religion, Scientificism

Introdução

Muito embora a ciência não seja inicialmente um sistema de crença tradicional, ao longo dos anos observou-se que crenças não religiosas podem ocasionar fenômenos geralmente associados a religião como a fonte de conforto, significado e demais funções compensatórias (Sagan, 2006; Rutjens et al., 2009; Rutjens et al., 2010b; Faria et al., 2013). Desta forma, a hipótese de Faria et al. (2013) é que a crença no valor da ciência como instituição e como uma das muitas formas de conhecimento pode apresentar o mesmo papel reconfortante que a crença secular em religião apresenta. O objetivo do instrumento *BISS (Believe in Science Scale)* é medir a crença do indivíduo na ciência, com questões autodeclaradas acerca da concordância/discordância e obviamente, acerca do quanto a ciência pode responder sobre questões alvo ou corroborar sua visão de mundo.

O debate sobre ciência enquanto crença, seja no sentido de refutar ou corroborar outras crenças, não é um produto novo e alguns dos registros citados por Machado et al. (2019) permanecem observáveis também nas novas mídias de comunicação. Nestes veículos, a divulgação científica mantém os objetivos sintetizados por Fuller (1998) de trazer produções científicas adaptadas a uma linguagem mais acessível e palatável a públicos diversos e incentivar que os indivíduos conheçam mais sobre ciência.

Assim como a crença, Rutjens et al. (2018) apresenta evidências da descrença científica, muitas vezes relacionadas a crenças religiosas ou até mesmo a visões políticas mais conservadoras (Lewandowsky & Oberauer, 2016) ou a ideias conspiracionistas diversas (Lewandowsky et al., 2013). De uma forma geral, há consideráveis estudos e evidências de que o pensamento analítico esteja relacionado a descrença religiosa (Gervais & Norenzayan, 2012; Stagnaro et al., 2019) enquanto que a crença religiosa estaria mais relacionada a intuição e por sua vez a processos automáticos e heurísticos.

Este estudo apresenta uma versão da adaptação do instrumento *BISS (Believe in Science Scale)* ao contexto brasileiro, bem como a avaliação dos indícios de validade transcultural, validade discriminante, validade de construto, de conteúdo e de critério da Escala de Crença em Ciência. O estudo objetivou ainda avaliar a relação da crença com

outros construtos e fatores como religião, grau de instrução e sexo, além de comparar médias de crença entre diferentes grupos religiosos.

Referencial Teórico

A Crença na Ciência

Além da relação entre razão e intuição no processo decisório, existem evidências de que traços de personalidade também possuam correlação significativa com religiosidade e crença (Martin & Nichols 1962; Kaldestad, 1992; Homayouni, 2011; Khoynzhad et al, 2012). As observações de Faria et al. (2013) e Shenhav e Greene (2011) evidenciam que quanto maior o estímulo a mentalidade intuitiva, maior a crença religiosa; enquanto que o pensamento analítico tende a aumentar a descrença religiosa ou mágica. Se de um lado há a crença, há uma contrapartida de descrença. Algumas evidências sobre as origens da descrença científica observadas por Rynes et al. (2018) podem ser ocasionadas por erros analíticos ou até mesmo por práticas de pesquisa questionáveis com falsificação de dados ou resultados. Outro ponto em relação a descrença pode ter correlação com a polarização ideológica que leva os indivíduos a buscar aconselhamento com pessoas que possuem o mesmo direcionamento de crenças (Mark et al., 2018).

Tais constatações não significam que crenças diversas não possam coexistir em um mesmo indivíduo e tampouco que estas crenças sejam contraditórias em sua totalidade, de maneira que um indivíduo pode costumeiramente apresentar múltiplas crenças sem que isso traga conflitos internos. Fatos como o exposto foram observados por Irwin et al. (2015) sobre um suposto pensamento duplo. O pesquisador propõe que essa modalidade de pensamento duplo (crer em ciência e crer no paranormal) ocorreria de forma compartimentada, mantendo ambas as formas de pensamento ao mesmo tempo. Fatores que poderiam explicar a tentativa de redução de uma possível dissonância seriam a busca seletiva (confirmatória ou para refutação) e a superestimação e/ou subestimação de informações de maneira a reduzir a distância entre ambas, como proposto por Schünemann (2012). Outro exemplo neste contexto é referente ao estudo de Machado e Zangari (2016) em que espíritas kardecistas e católicos utilizaram-se de argumentação racional e com base científica para justificar

fenômenos e alinhar crenças. Outra semelhança com esta situação foi discutida por Krauss e Colombo (2016) em que foi sugerido que evidências ambíguas possam ser interpretadas como apoiadoras de uma conclusão.

Letramento Científico Brasileiro

A discussão sobre a possibilidade de que a educação científica possa ou não reduzir a crença em fenômenos paranormais foi realizada no estudo de Maraldi et al. (2016) e, embora existam sugestões de que a crença no paranormal seja reduzida conforme se eleva o letramento científico, apresentam-se também evidências de que não existe relação entre a crença no paranormal e o letramento científico (Broch, 2000).

Os índices de conhecimento formal e, por consequência, letramento científico são baixos no Brasil. De acordo com o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP, 2017), em relação aos índices da escala do Sistema de Avaliação da Educação Básica [Saeb] de aprendizado que possui nove níveis, a média de proficiência em língua portuguesa e em matemática ficou em 210 (português) e 220 (matemática), o que equivale ao 4º nível. Os dados fornecidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2020) demonstram que em 2015 quase 56% da população possuía menos de 10 anos de estudo, sendo que 7% da população foi considerada como analfabeta. Tais números contrastam com o Indicador de Analfabetismo Funcional - Inaf do Instituto Paulo Montenegro que apresentam sensíveis 29% da população como o número de indivíduos considerados como analfabetos funcionais em diferentes níveis na população brasileira. No ensino superior, 4% da população foi considerada como analfabetos funcionais na pesquisa daquele ano. Vale observar também os resultados de indicadores internacionais como o do Programa Internacional de Avaliação de Estudantes [PISA]. Na edição de 2018, foram avaliados 79 países e o Brasil figura entre as 15 últimas posições do *ranking* em desempenho escolar em leitura, matemática e ciências.

O Brasileiro e a Crença na Ciência

No presente estudo, realizou-se a adaptação do instrumento de Crença em Ciência para o contexto brasileiro e buscou-se identificar os indícios de validade interna

(convergente, discriminante, conteúdo e transcultural) e validade externa derivadas do conceito de rede nomológica (Cronbach & Meehl, 1955; Pasquali, 2007; Pais-Ribeiro, 2013), considerando que a crença na ciência se relacione com a religiosidade, com traços de personalidade, racionalidade e sexo biológico. Já existe um estudo com a utilização de uma versão da escala em que são relatados os resultados da análise fatorial exploratória como indicador de qualidade psicométrica, além do índice de concordância entre os juízes. Este manuscrito não trata de uma replicação do estudo de Machado et al. (2019), e direciona-se a discutir os indícios de validade observados com os resultados do processo de adaptação, julgamento por juízes e especialistas, análise fatorial exploratória, análise fatorial confirmatória, análise da correlação entre as variáveis e comparação entre as médias de grupos.

Método

A Adaptação ao Contexto Brasileiro

A importância da adaptação e de procedimentos metodológicos adicionais, muito mais complexa do que a operacionalização literal de uma tradução, é enfatizada por Sireci, (2005) e Borsa et al. (2012), cuja orientação versa sobre evitar a destruição da validade do construto e promover a manutenção do conceito atribuído ao instrumento. Então, optou-se por conduzir um novo procedimento de tradução como uma das maneiras de observar evidências de que o conteúdo do item tenha se mantido e seja adequado ao contexto brasileiro.

Portanto, alguns procedimentos orientados por Borsa et al. (2012) envolvem tradução e tradução reversa, síntese de versões, avaliação por juízes, avaliação pelo público-alvo, estudo piloto e a avaliação da estrutura fatorial do instrumento e desta forma auxiliariam ao pesquisador na tentativa de manter o contexto e observar indícios de validade. A tradução inicial foi realizada por profissional experiente e com formação de Bacharelado em Letras e Tradução, além de traduções livres realizadas por outros indivíduos com notório saber na língua inglesa. Desta forma, as propostas de utilização de

múltiplas versões com tradutores bilíngues e sem familiaridade com o construto foram atendidas (Beaton et al., 2000; Cassep-Borges et al., 2010).

Para testar a consistência em uma situação de tradução reversa (Beaton et al., 2000; Sireci, 2005; Borsa et al., 2012), a versão de consenso foi submetida a dois universitários nativos em língua inglesa residentes nos Estados Unidos que possuem o português como segunda língua. O contexto neste processo foi mantido e apenas alguns termos tiveram alteração, sem prejuízo ao significado. Interessa também comparar a versão utilizada por Machado et al., (2019) com a versão de consenso deste estudo, conforme se observa na tabela 1:

Tabela 1

Comparação da versão final da pesquisa com outra versão já utilizada

Versão da pesquisa	Tradução
1. A ciência nos fornece uma melhor compreensão do universo do que a religião.	1. A ciência nos proporciona uma melhor compreensão do universo do que a religião.
2. Em um mundo assombrado por demônios, a ciência é uma vela no escuro. (Carl Sagan)	2. “Num mundo com tantas superstições e cheio de credices/crenças infundadas, a ciência é como uma luz no fim do túnel.” (Carl Sagan).
3. Só podemos acreditar racionalmente no que é cientificamente comprovável.	3. Só podemos acreditar de forma racional no que é cientificamente comprovável.
4. A ciência nos diz tudo o que há para saber sobre a realidade.	4. A ciência é suficiente para explicar o que é a realidade.
5. Todos os obstáculos que os humanos enfrentam podem ser resolvidos com a ciência.	5. Todas as tarefas com as quais os seres humanos se deparam podem ser resolvidas pela ciência.
6. O método científico é o único caminho confiável para o conhecimento.	6. O método científico é o único caminho confiável que leva ao conhecimento.
7. O único tipo real de conhecimento que podemos ter é o conhecimento científico.	7. O conhecimento científico é o único tipo de conhecimento verdadeiro que podemos ter.
8. A ciência é a parte mais valiosa da cultura humana.	8. A ciência é o que há de mais valioso na cultura humana.
9. A ciência é o meio mais eficiente de alcançar a verdade.	9. A ciência é o meio mais eficiente de alcançar a verdade.
10. Os cientistas e a ciência devem ter mais respeito na sociedade moderna.	10. Os cientistas e a ciência deveriam ser mais respeitados em nossa sociedade.

Ao analisar a tabela 1, no que tange a convergência, é possível observar que os itens 6 e 9 apresentam a integralmente a mesma tradução enquanto que os demais itens apresentam alguma diferença de palavras (por exemplo de “racionalmente” para “de forma racional”, sem que se altere o contexto). Em relação ao item 2, a tradução realizada pelos pesquisadores suprimiu o contexto metafórico de Sagan (2006), mas ainda assim manteve o contexto importante presente na afirmação, em que a ciência é vista como uma forma esperançosa de resolução de um problema (a luz no fim do túnel ou a vela na escuridão).

O grupo de juizes foi composto por 3 Bacharéis em Administração de empresas e 1 Bacharel em Economia, todos mestrandos em Psicologia do Trabalho, 3 Doutorandos em Psicologia e um Doutor em Psicologia. A importância da quantidade de pessoas envolvidas no julgamento e das notas atribuídas a qualidade do conteúdo são sugeridas por Alexandre e Coluci (2011) para que sejam observadas a relevância, a clareza e a pertinência dos itens. Os referidos autores assinalam um índice de concordância acima dos 90% entre os juizes como critério para atribuir validade de conteúdo. Os especialistas e juizes foram convidados a atribuir notas em uma escala de 1 a 5, considerando que 5 representaria uma concordância muito alta em relação ao quesito e 1 representaria uma concordância muito baixa. A partir daí, calculou-se o Índice de Validade do Conteúdo que foi de 93%, considerando a quantidade de notas 4 ou 5 em relação ao total de itens e de respostas.

Variáveis Presentes no Manuscrito

Nesta pesquisa foram utilizadas as variáveis religião (transcrita como a afiliação religiosa que o indivíduo declara pertencer, apresentando algumas das principais denominações) o sexo, o nível de escolaridade, a idade, região demográfica, traços de personalidade - aqui baseado no instrumento IGFP-5 que trata de uma adaptação e apresentação de indícios de validade *Big Five Inventory* foi realizada por Andrade (2008) que é composto por 5 fatores (conscienciosidade, abertura à experiência, amabilidade, extroversão e neuroticismo) – e o instrumento chamado *Cognitive Reflection Test* que apresenta uma medida de razão e intuição.

Participantes e Pressupostos Estatísticos

A coleta ocorreu entre os anos de 2019 e 2020 em uma pesquisa intitulada “o que você pensa sobre ciência?” totalizando 662 observações. Dentre as informações coletadas estão dados sociodemográficos, a aplicação do instrumento de crença em ciência, uma versão do IGFP-5 (*BIG FIVE*) e um instrumento que mede racionalidade/intuição (CRT). No campo religião, foram adotados os agrupamentos acima de 20 indivíduos. As religiões com menos de 20 respondentes foram agrupadas na classe “outras”. Os dados agrupados podem ser observados na tabela 2.

Tabela 2*Caracterização dos participantes da pesquisa*

Características	Distribuição	
	<i>n</i>	%
Sexo	Homem	293 44%
	Mulher	369 56%
Escolaridade	Nível Médio	75 11%
	Superior-Graduação	207 31%
	Superior -Pós-Graduação -Especialização	332 50%
	Superior – Mestrado	39 6%
	Superior – Doutorado	9 1%
Idade	18-25	40 6%
	26-30	57 9%
	31-35	121 18%
	36-40	175 26%
	41-45	107 16%
	46-50	68 10%
	51-55	45 7%
	56-60	27 4%
	61-65	14 2%
	66-70	6 1%
71-75	2 0%	
Região	Centro Oeste	162 24%
	Nordeste	34 5%
	Norte	13 2%
	Sudeste	399 60%
	Sul	54 8%
Religião	Ateus/Agnósticos	125 19%
	Católicos	228 34%
	Evangélicos/Protestantes	80 12%
	Kardecistas	80 12%
	Cristão – Outros	44 7%
	Matrizes africanas	26 4%
	Outras	79 12%

A não normalidade dos dados poderia ser um fator de alerta caso a amostra fosse pequena e com existência de heterocedasticidade (Hair et al, 2009; Ghasemi & Zahediasl, 2012; Pino, 2014; Chamtarangsi et al., 2016; Miot, 2017). Para esta amostra, foram adotados os procedimentos para prosseguir a análise, visto o tamanho ($n > 200$ indivíduos) e indícios de homocedasticidade. Outra observação importante para corroborar a decisão é que a assimetria fica acima de 1 e abaixo de 2 somente no item 10, o que segundo Miles & Shevlin (2001) poderia ser considerado para o prosseguimento das análises, uma vez que tal fato não seria um indicativo de grandes problemas quanto ao viés. Também, não foram encontrados valores omissos na amostra.

Pela análise do teste M de Box, foi observada a não rejeição da hipótese nula, visto a significância considerando $p < 0,05$, o que indica igualdade de variância entre os itens.

Resultados

Os resultados das análises objetivam subsidiar a discussão sobre os indícios de validade do instrumento. Sobre a validade interna, a análise fatorial confirmatória e exploratória serão consideradas para observação de indícios de validade fatorial, convergente (confiabilidade composta e variância média extraída) e discriminante. Sobre a validade externa e de critério, serão realizadas análises das correlações e a comparação da diferença entre grupos, sob a ótica da rede nomológica.

Estudo 1 - Análise Fatorial Exploratória e Confirmatória

Para a realização da Análise Fatorial Exploratória, foi utilizada uma amostra com 662 indivíduos e os fatores foram extraídos pelo executável *Factor* versão 10.10.03, utilizando os métodos *Robust Diagonally Weighted Least Squares* e *Minimum Rank Factor Analysis* via análise paralela (Damásio, 2012; Machado et al., 2014; Hernandez, 2015; Gomes et al., 2016) com rotação Promax e supressão de valores abaixo de 0,30, sem a indicação do número de fatores a extrair. Nas ciências humanas e sociais, geralmente são utilizadas as rotações oblíquas, visto que conceitualmente se pressupõe relação entre os fatores, fato este atrelado ao cenário de que é impossível isolar os fatores adequada e integralmente (Damásio, 2012; Tabachinick & Fidel, 2014).

A fatorabilidade da matriz é observada pelo teste de Kaiser-Meyer-Olkin ($KMO=0,92$) e o teste de Esfericidade de Bartlett (significante a $X^2=4252,6; 45; p>0,000$). Ambos índices sinalizam boas evidências neste estudo, visto a adequação tanto ao pressuposto de que índices de KMO superiores a 0,7 indicam boa fatorabilidade (Hutcheson & Sofroniou, 1999; Laros, 2004; Damásio, 2012) e a detectada significância no teste de esfericidade (Tabachinick & Fidel, 2014). Os fatores extraídos apresentaram cargas superiores a 0,5 (Hair et al., 2009) e comunalidades acima de 0,3 e inferiores a 0,8 (Field, 2018), pressupostos estes observados na análise. A proporção de variância explicada pelo instrumento foi de 0,64 com a indicação de apenas um fator.

Observa-se um Alfa de Cronbach elevado ($\alpha=0,96$) assim como o Lambda de Guttman ($\lambda^2=0,97$), evidências estas de que o instrumento possui excelentes índices de

confiabilidade (Maroco & Garcia-Marques, 2013; Souza et al,2017). A mesma qualidade de índice de confiabilidade foi observada em cada uma das questões de maneira individual. Desta forma, torna-se desnecessária a exclusão de itens com o objetivo de melhoria nos níveis de confiabilidade.

Com o intuito de realizar a avaliação de índices de validade discriminante, foram obtidas as correlações ao quadrado dos itens que compõem a escala para comparação com a variância média extraída. A média das correlações ao quadrado observada na análise foi de 0,40. Embora as cargas fatoriais em sua maioria apresentem-se elevadas e acima de 0,50, seus demais índices (relação $X^2/$ Graus de liberdade, CFI, TLI, RMSEA e VME) transcritos na tabela 3 como modelo 1 não apresentaram bons resultados sob a ótica dos pressupostos da análise fatorial confirmatória.

Tabela 3

Relação dos demais índices da análise

	<i>Observado</i>	<i>Esperado</i>
Qui Quadrado	385,35	-
GI	35	-
P-valor	0,000	<0,05
X^2/GL	11,01	<5
CFI	0,98	>0,9
GFI	1,00	>0,9
RMSEA	0,12	<0,08
FDI	0,97	>0,90
CC	0,94	0,70
VME	0,61	0,50
A	0,96	>0,70

Nota: índices esperados conforme Brown (2006) e Hair et al (2009)

Conforme se observa, a exclusão de itens traz evidências de que há melhoria nos índices apresentados na análise. Dentre os indicadores, apenas o RMSEA não estaria enquadrado nos parâmetros esperados. Mas é possível observar que, apesar de marginalmente localizado em seu ponto médio, o limite inferior do RMSEA é de 0,08 e cruza a faixa de valor indicado. A discussão acerca dos limites para aceitação dos índices de ajuste em diversos artigos (Rigdon, 1996; Hu & Bentker, 1999; Lai & Green,2016; Lai, 2019; Xia &Yang,2019) convergem quanto a utilização conjunta de índices como o RMSEA, TLI, GFI e CFI de maneira a complementar o processo de decisão para aceitação do modelo.

Como os demais índices indicam o ajuste, é possível corroborar a hipótese de bom ajuste do modelo, visto a proximidade do RMSEA ao valor esperado.

Estudo 2 – Crença na Ciência e Suas Correlações

Com o intuito de observar as relações entre a crença na ciência e traços de personalidade (Martin & Nichols 1960; Kaldestad, 1992; Homayouni, 2011; Khoynzhad et al, 2012), pensamento analítico (Gervais & Norenzayan, 2012; Sanchez et al, 2017), religiosidade (Rutjens et al.,2018), sexo e grau de instrução (Maraldi et al.,2016), foi realizada uma análise das correlações considerando as variáveis observadas na Tabela 4. As observações foram elencadas de maneira a demonstrar a relação da crença na ciência com alguns dos fatores e construtos relacionados na literatura e presentes neste estudo.

Tabela 4

Correlação de Spearman da crença com outros fatores

	<i>Índices</i>	
	<i>Correlação</i>	<i>p>0,05</i>
CRT	0,051	0,189
Abertura	0,157**	0,000
Amabilidade	-0,227	0,000
Conscienciosidade	-0,159**	0,000
Extroversão	-0,107	0,006
Neuroticismo	-0,017	0,694
Religião	-0,450**	0,000
Idade	-0,103	0,009
Região demográfica	0,073	0,070
Sexo	-0,237**	0,000
Escolaridade	-0,040	0,310
Evangélicos	-0,297	0,000
Ateu e Agnósticos	0,450	0,000
Católicos	-0,103	0,008
Kardecistas	0,021	0,582

Nota: ** significante a $p>0,01$

As correlações observadas são consideradas de fraca (abertura de maneira positiva e idade, sexo, amabilidade e conscienciosidade de maneira negativa) a moderadas (grupo ateu/agnósticos como positiva e grupo evangélico e religião como negativa). A variável religião é composta da totalidade de manifestações religiosas presentes na pesquisa, enquanto que a variável considerando o grupamento religioso foi incluída na tabela para que fosse possível mensurar a sua relação com a crença na ciência de maneira individualizada de algumas das religiões. A escolha por individualizar a observação da relação de algumas das denominações religiosas foi baseada na quantidade de respondentes presentes no estudo, além do quantitativo e representatividade nacional das religiões observada no último

censo realizado em 2010, considerando as 4 maiores representações. Embora algumas das relações apresentem-se de maneira quase irrisória, foram incluídas para efeito de avaliar algumas das hipóteses sugeridas na literatura, que serão discutidas posteriormente.

Estudo 3 – Grupos Religiosos e a Crença na Ciência

Neste estudo, além de agrupar as denominações religiosas de acordo com sua similaridade conceitual, foi realizado também o nivelamento da quantidade de indivíduos por grupo religioso, de maneira a comparar as médias entre ateus, evangélicos, católicos e kardecistas. A amostra foi separada com o objetivo de garantir que cada grupo deste estudo tivesse no mínimo 80 indivíduos. Foram selecionados 4 grupos como ateus, católicos, evangélicos e kardecistas, agrupamentos estes que compõem grande parte das observações de denominação religiosa desta pesquisa, além constarem como as 4 maiores representações observadas no último censo do IBGE de 2010.

Por meio de uma ANOVA, [$Z(3,316) = 59,73$; $p < 0,05$], considerando o teste robusto de igualdade das médias, foi rejeitada a hipótese nula - H_0 de que os grupos possuem igualdade de médias. Na comparação *Post Hoc* o que se observa é que existe igualdade somente na comparação entre dois grupos (espíritas e católicos), o que pode ser observado na Tabela 5. O que se observa é que as médias são matematicamente diferentes e apresentam diferenças estatisticamente significantes na comparação entre o grupamento de kardecistas e católicos (que possuem médias estatisticamente iguais) em relação aos grupos ateu/agnósticos ou evangélicos.

Tabela 5

Comparação entre grupos via Teste Tukey

<i>Médias e subconjuntos</i>			
	1	2	3
Evangélicos	2,878		
Kardecistas		3,548	
Católicos		3,648	
Ateus			4,859
Sig	1,000	0,911	1,000

Nota: considerando $n=80$ para cada grupo

Embora os pressupostos de normalidade e homogeneidade de variâncias apresentem-se violados, há indícios de que os impactos sejam minimizados, visto que se

trata de uma amostra com mais de 200 indivíduos (Hair et al, 2009) e que apresentou bons índices de tamanho do efeito (η^2 parcial quadrado considerável de 0,36) e um poder de 0,999 calculado via *G*power* 3.1.9.4.

Discussão

Esta pesquisa apresenta indícios de validade interna e externa para a versão brasileira da escala de Crença em Ciência.

Inicialmente, a pesquisa tratou de apresentar uma versão do instrumento BISS (Crença em Ciência) e os consequentes indícios da adequação dessa versão aos requisitos sobre a adaptação do instrumento (Beaton et al., 2000; Sireci et al., 2005; Borsa et al., 2012; Cassep-Borges et al., 2010), o que corrobora a hipótese das boas qualidades psicométricas do instrumento. O processo de adaptação com itens próximos em termos de conteúdo ao original e o índice de validade de conteúdo acima dos 90% são indícios de validade transcultural e de conteúdo, antecedentes dos bons resultados observados na pesquisa.

No que tange aos indícios de validade interna e mais precisamente considerando a análise fatorial exploratória realizada no estudo 1, a estrutura fatorial encontrada teve bons índices de fatorabilidade e uma considerável variância explicada de 64%, além de excelente confiabilidade observada em seu Alfa de Cronbach ($\alpha=0,96$) convergente com estudos internacionais (Farias et al., 2013; Irvin et al., 2015; Stal et al, 2016; Dagnal et al., 2019, Machado et al., 2019). Os itens apresentaram cargas fatoriais elevadas, corroborando resultados observados no estudo de Dagnal et al. (2019), bem como os demais índices do instrumento que sinalizam a aderência a utilização da técnica utilizada.

Os indícios de validade convergente estimados pela Variância Média Extraída (VME) e pela Confiabilidade Composta (CC) (Valentini & Damásio, 2016) também foram observados no modelo. No que tange a validade discriminante o pressuposto referido diz respeito a VME observada, que foi maior do que as correlações dos itens ao quadrado. Como não há outros fatores, a comparação foi feita entre o VME e o quadrado da correlação entre os itens. Os demais índices também apresentaram-se satisfatórios e se adequaram aos valores considerados para bons ajustes (Kline, 2005; Hair et al 2009).

No que tange aos indícios de validade externa, as evidências se apresentam pelas correlações encontradas com outros construtos e fatores elencados na literatura e que foram objeto de análise nesta pesquisa. Baseado nos conceitos da rede nomológica construídos ao longo dos anos sobre a hipótese de que a validade convergente e discriminante é mensurável pelo nível em que um construto se relaciona ou difere de outros construtos, fatores ou critérios (Crombach & Meehl, 1955; Campbell & Fiske, 1959; Brown, 1970; Pasquali, 2007; Pais-Ribeiro, 2013; Viana, 2015). São considerados como indícios de validade convergente, discriminante e critério as relações que o construto em questão apresentou nos testes realizados com as variáveis propostas. Tais relações são discutidas nesta sessão de maneira a comparar a referência na literatura com aquilo que se observou na pesquisa.

As relações observadas no estudo 2 corroboram que exista uma relação fraca entre a personalidade (considerando os traços de personalidade inseridos na teoria dos cinco grandes fatores) e a crença. Também, a crença ou descrença religiosa (comparando-se aqueles que acreditam em alguma religião em contrapartida com aqueles que são ateus ou agnósticos) não apresentou correlação com a racionalidade, sugerindo que o pensamento analítico não reduziu a crença religiosa (Gervais & Norenzayan, 2012; Sanchez et al, 2017). Nesse manuscrito, a medida de racionalidade utilizada (CRT) não apresentou uma correlação significativa com a crença na ciência, diferente das observações de outros estudos que apontam uma relação entre a crença na ciência e a racionalidade.

Sobre os traços de personalidade representados neste estudo pelo instrumento IGFP-5 (*Big Five*), destacam-se as observações atreladas aos fatores conscienciosidade, amabilidade e abertura à experiência como as mais relevantes. A abertura à experiência está relacionada ao nível de curiosidade e complexidade de pensar, a conscienciosidade está relacionada a perseverança e a busca por objetivos (Pervin & John, 2004) e a amabilidade está relacionada ao nível de socialização e antagonismo com os demais indivíduos (Hutz et al, 1998). Algumas das relações observadas se apresentam semelhantes as definições de ciência sugeridas por Popper (1972), baseada na busca complexa, criativa

e pouco emocional (antagônica e mais racional) do entendimento do mundo, porém um pouco diferentes do esperado no que tange a conscienciosidade, fator este muito observado na literatura como preditor de desempenho acadêmico (Nechita et al, 2015) e bem-estar pessoal nas organizações (Dessen & Paz, 2010).

Observou-se também uma relação significativa entre o sexo e a crença científica, de modo os escores apresentados pelos homens obtiveram relação positiva com maiores escores de crença na ciência quando comparados ao escores apresentados pelas mulheres. Outro fato observado e de importante consideração é que o grau de escolaridade não apresentou nenhuma relação com a crença científica, refutando a hipótese de que existisse uma correlação entre o quantitativo de anos de estudo e a maior crença na ciência, em consonância com o que foi discutido por Maraldi et al. (2016). As evidências observadas por Irwin et al., (2016) e Machado et al (2019) são de que, no meio acadêmico, a religiosidade convive com o método científico, de maneira que aqueles que possuem uma afiliação tradicional (dentre as costumeiramente encontradas no Brasil), muitas vezes separam a racionalidade por trás da ciência de suas crenças pessoais, e este fato é corroborado pelos achados desta pesquisa.

A correlação com a crença ou descrença religiosa também sugere que visões religiosas menos fundamentalistas ou irreligiosas se relacionem positivamente com a crença na ciência. O ateísmo e agnosticismo, por exemplo, apresentaram as maiores associações ($r=0,45$). Tais achados no contexto brasileiro corroboram estudos internacionais sobre a natureza da crença e algumas de suas relações.

No estudo 3, foram comparadas as médias da crença entre os principais grupos religiosos e observou-se que embora exista uma diferença matemática e passível de ordenação entre católicos, espíritas kardecistas e evangélicos, tal diferenciação não foi estatisticamente significativa. A diferença significativa deu-se na comparação com o grupo de ateus e agnósticos, que apresentaram a maior média de crença. O poder e o tamanho do efeito, observados de maneira satisfatória, auxiliariam na comparação com estudos futuros.

Em diversas áreas, a crença na ciência ainda é um campo com poucos estudos. No mundo das organizações e do trabalho, são escassos os estudos sobre o letramento científico dos trabalhadores e a baixa crença científica ainda não se apresenta como um problema notável, embora a descrença na ciência e no processo racional abra espaço para uma gestão mais temerária que contraria abordagens de gestão baseadas em evidências e premissas organizacionais atuais. Tal cenário de descrença ou desconhecimento científico também já foi observado por Cruz et al. (2008) que comentam sobre a constância de teorias sem embasamento científico costumeiramente utilizadas nas organizações. A proliferação de teorias sem utilização de método científico nas organizações movimentam cifras elevadas e se apresenta como mais um dos muitos nichos de mercado consolidados.

Algumas das limitações deste trabalho estão vinculadas ao fato de que não foi utilizada uma escala para medir o grau de religiosidade dos participantes e que talvez justificasse eventuais correlações positivas ou negativas de determinadas crenças religiosas com a crença em ciência em certos níveis. Não foi possível observar de maneira escalar em que momento os níveis de crença religiosa e científica começam a convergir ou divergir. Uma medida de coabitação de ambas as crenças pode ser explorada em estudos futuros. O sincretismo religioso também é ponto de observação em pesquisas desta natureza, visto que é comum encontrar indivíduos que se declaram como adeptos de uma religião, mas que realizam práticas que seriam atribuídas a outras crenças, tornando complexo classificar de maneira simplista a crença religiosa de um indivíduo. Outro ponto importante diz respeito a representatividade de algumas religiões que, embora possam parecer semelhantes e muitas vezes agrupáveis em um nível macro, têm visões muito distintas se analisadas de maneira mais pormenorizada. Mesmo dentro de uma mesma denominação religiosa, como por exemplo os católicos ou evangélicos, há diversas subdivisões em relação a devoção, agrupamento em comunidade ou ordem. Então, estudos mais avançados podem ainda apresentar melhor a taxonomia e classificação das manifestações religiosas em relação a suas crenças não religiosas.

Importante ressaltar que a crença na ciência não implica necessariamente em letramento científico ou conhecimento de fato sobre como o método científico é conduzido, de modo que não foram utilizadas medidas para avaliar o nível de conhecimento ou familiaridade com termos correlatos a ciência que os indivíduos pesquisados possuíam. Uma hipótese ainda em aberto é que embora a quantidade de anos de estudo tenha se mostrado não significativo quanto a sua relação com a crença na ciência, é possível que a qualidade do ensino possa apresentar alguma relação como o nível de crença científica.

Assim, conclui-se que as evidências apontam para boas qualidades psicométricas do instrumento e que o construto da crença na ciência pode ser medido de maneira adequada para o contexto brasileiro. Tal possibilidade amplia horizontes de pesquisa em que esse tipo de variável é relevante e ainda pouco explorada, como por exemplo, nas organizações. Os bons índices encontrados neste estudo apontam uma luz sobre a utilização no contexto brasileiro.

Referências

- Amorim, L. (2015). Análise fatorial confirmatória e modelos com equações estruturais: Um Tutorial usando Software Estatístico. 10.13140/RG.2.1.1520.4960.
- Andrade, J. M. (2008). *Evidências de validade do Inventário dos Cinco Grandes Fatores de personalidade para o Brasil*. Tese de Doutorado, Universidade de Brasília.
- Andrade, M. O. (2019), A religiosidade brasileira: o pluralismo religioso, a diversidade de crenças e o processo sincrético. *CAOS – Revista Eletrônica de Ciências Sociais*, v. 2, n. 14, p. 106-118.
- Barends, E., Janssen, B., Have, W., & Have, S. (2014). Difficult but doable: Increasing the internal validity of organizational change management studies. *Journal of Applied Behavioral Science*, 20(5),1-5. doi: 10.1177/0021886313515614
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186-3191.

- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Adaptação e Validação de Instrumentos Psicológicos entre Culturas: Algumas Considerações. *Paidéia*, 22(53), 423–432. <https://doi.org/10.1590/S0103-863X2012000300014>
- Bortolotti, R. G. (2014). Religious belief in peirce's semiotic perspective/A crenca religiosa sob a perspectiva semiotica de peirce. *Acta Scientiarum. Human and Social Sciences*, 36(2), 235+.
- Broch, H. (2000). Save our science: the struggle for reason at the university. *Skeptical Inquirer*, 24 (3), 34-39
- Brown, F. G. (1970) *Principles of Educational and Psychological testing*. The Dryeden Press Inc. Hinsdale. Illiniois
- Brown, T. A (2006); *Confirmatory fator analysis for Apllied Research*, 1st Edition, New York, The guilford.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105
- Cassepp-Borges, V., Balbinotti, M. A. A., & Teodoro, M. L. M. (2010). Tradução e validação de conteúdo: Uma proposta para a adaptação de instrumentos. In L. Pasquali, *Instrumentação psicológica: Fundamentos e práticas* (pp. 506-520). Porto Alegre: Artmed.
- Chantarangsi, W., Liu, W., Bretz, F., Kiatsupaibul, S., & Hayter, A.. (2016). Normal Probability Plots with Confidence for the Residuals in Linear Regression. *Communications in Statistics - Simulation and Computation*. 47. 10.1080/03610918.2016.1165840.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302
- Cruz, E., Fabio, Siqueira, N & Silva, Da. (2008). Ciência e Pseudociência na Administração. *Revista Pensamento Contemporâneo em Administração*. 2. 10.12712/rpca.v2i2.131.
- Dagnall, N., Denovan, A., Drinkwater, K. G., & Parker, A. (2019). An Evaluation of the Belief in Science Scale. *Frontiers in psychology*, 10, 861. <https://doi.org/10.3389/fpsyg.2019.00861>

- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, 11(2), 213-228.
- Dessen, M. C., & Paz, M. G. T. (2010). Bem-estar pessoal nas organizações: o impacto de configurações de poder e características de personalidade. *Psicologia: Teoria e Pesquisa*, 26(3), 549-556. <https://doi.org/10.1590/S0102-37722010000300018>
- Farias, M. Reiman, A. K., Kahane, G., & Toledo, Z. (2013). Scientific faith: Belief in science increases in the face of stress and existential anxiety. *Journal of experimental social psychology*. 49. 1210-1213. 10.1016/j.jesp.2013.05.008.
- Fuller, G. (1998) *Cultivating science: negotiating discourse in the popular texts of Stephen Jay Gould*. In: Martin, J., Veil R. (Ed.). *Reading science: critical and functional perspectives on discourses of science*. London and New York: *Routledge*,. p. 35-62.
- Gervais, W. M., & Norenzayan, A. (2012) Analytic thinking promotes religious disbelief. *Science*. Apr 27;336(6080):493-6. doi: 10.1126/science.1215647.
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Hair, Jr., J.F.; Willian, B.; Babin, B. & Anderson, R. E. (2009). *Análise multivariada de dados*. 6.ed. Porto Alegre: Bookman.
- Homayouni, A. (2011). The role of personality traits and religious beliefs in tendency to addiction. *Procedia - Social and Behavioral Sciences*, 30, 851–855. doi:10.1016/j.sbspro.2011.10.165
- Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. London: Sage Publications.
- Hutz, C. S., Nunes, C.H., Silveira, A. D., Serra, J., Anton, M., & Wieczorek, L. S. (1998). O desenvolvimento de marcadores para a avaliação da personalidade no modelo dos cinco grandes fatores. *Psicologia: Reflexão e Crítica*, 11(2), 395-411. <https://doi.org/10.1590/S0102-79721998000200015>

Instituto Brasileiro de Geografia e Estatística (2015). *Distribuição das pessoas de 25 anos ou mais de idade, por grupos de estudo – Brasil 2007/2015.*

<https://brasilemsintese.ibge.gov.br/educacao/anos-de-estudo.html>

Instituto Brasileiro de Geografia e Estatística (2015). *Taxa de analfabetismo das pessoas de 10 anos ou mais de idade, por sexo – Brasil 2007/2015.*

<https://brasilemsintese.ibge.gov.br/educacao/taxa-de-analfabetismo-das-pessoas-de-10-anos-ou-mais.html>

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2017).

Apresentação e Resultados das aplicações amostrais para 2º e 9º ano do ensino fundamental. <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb/resultados>

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2020). *Programa Internacional de Avaliações de Estudantes (PISA).* <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/pisa>

Instituto Paulo Montenegro (2015). *Indicador de Alfabetismo Funcional – INAF- Estudo especial sobre analfabetismo e mundo do trabalho.*

<https://drive.google.com/file/d/0B5WoZxXFQTCRRWFyakMxOTNyb1k/view>

Instituto Paulo Montenegro (2018). *Indicador de Alfabetismo Funcional INAF Brasil 2018.*

<https://drive.google.com/file/d/1ez-6jrlrRRUm9JJ3MkwxEUffltjCTE16/view>

Irwin H. J., Marks A. D. (2013). The “Survey of scientifically unaccepted beliefs”: a new measure of paranormal and related beliefs. *Aus. J. Parapsychol.* 13 133–167

Irwin, H. J., Dagnall, N., & Drinkwater, K. (2015). The role of doublethink and other coping processes in paranormal and related beliefs. *J. Soc. Psych. Res.* 79, 80–97.

Irwin, H. J., Dagnall, N., & Drinkwater, K. (2016). Dispositional scepticism, attitudes to science, and belief in the paranormal. *Aus. J. Parapsychol.* 16, 117–131.

Irwin, H. J., Drinkwater, K., & Dagnall, N. (2015). The role of doublethink and other coping processes in paranormal and related beliefs. *Journal of the Society for Psychical Research*, 79(919[2])[2], 80–97.

- Kahneman, D (2012). *Rápido e devagar, duas formas de pensar*. Rio de Janeiro: Editora Objetiva.
- Kaldestad, E. (1992), "Religious orientation, personality, mental health and religious activity", *Nordic Journal of Psychiatry*, Vol. 36 No. 5, pp. 95-1
- Karadöller, D., Yılmaz, O. & Sofuoğlu, G. (2015). Analytic Thinking, Religion and Prejudice: An Experimental Testing of the Dual-Process Model of Mind. *The International Journal of Psychology of Religion* 10.13140/RG.2.2.24522.77764.
- Khoynezhad, G., Rajaei A. R., & Sarvarazemy A. (2012) Basic religious beliefs and personality traits. *Iran J Psychiatry. Spring;7(2):82-6..*
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling* (2a ed.). New York: The Guilford Press
- Krauss, A., & Colombo, M. (2020) Explaining public understanding of the concepts of climate change, nutrition, poverty and effective medical drugs: An international experimental survey. *PLoS ONE* 15(6): e0234036. <https://doi.org/10.1371/journal.pone.0234036>
- Lai, K. (2019). Creating Misspecified Models in Moment Structure Analysis. *Psychometrika*. 84. 10.1007/s11336-018-09655-0.
- Lai, K., & Green, S. B. (2016). The Problem with Having Two Watches: Assessment of Fit When RMSEA and CFI Disagree. *Multivariate behavioral research*, 51(2-3), 220–239. <https://doi.org/10.1080/00273171.2015.1134306>
- Laros, J. A. (2004). O uso da análise fatorial: algumas diretrizes para pesquisadores. Em: L. Pasquali (Org.), *Análise fatorial para pesquisadores*. Petrópolis: Vozes.
- Lewandowsky, S., Gignac, G.E., & Oberauer, K. (2013) The Role of Conspiracist Ideation and Worldviews in Predicting Rejection of Science. *PLoS ONE* 8(10): e75637. <https://doi.org/10.1371/journal.pone.0075637>
- Lewandowsky, S., & Oberauer, K. (2016). Motivated Rejection of Science. *Current Directions in Psychological Science*, 25(4), 217–222. <https://doi.org/10.1177/0963721416654436>

- Ives, J. E., Cavenaghi, S, Barros, L. F., & Carvalho, A. A., (2017). Distribuição espacial da transição religiosa no Brasil. *Tempo Social*, 29(2), 215-242. <https://doi.org/10.11606/0103-2070.ts.2017.112180>
- Machado, F. R., & Zangari, W. (2016). Omissões da psicologia, territórios colonizados: experiências anômalas, interpretações religiosas e atuação do psicólogo. In *Laicidade, religião, direitos humanos e políticas públicas*. São Paulo: Conselho Regional de Psicologia. Recuperado de <http://www.crsp.org/fotos/pdf-2016-06-21-18-16-42.pdf>
- Machado, F. R., Torres, C. M., Huang, M. F. C., Zangari, W., & Maraldi, E. O.(2019) . A Ciência dos Religiosos: Estudo Exploratório dos Usos e Sentidos que Religiosos Fazem da Ciência. *Interação em Psicologia (online)*, v. 23, p. 243-254.
- Magnavita, P. R. (2013). Ciência e religião: por uma micropolítica de subjetivação. *Caderno CRH*, 26(69), 493-510. <https://doi.org/10.1590/S0103-49792013000300006>
- Maraldi, E. O.; Machado, F. R. & Zangari, W. Pode (e deve) a educação científica diminuir a crença nos fenômenos paranormais? (2016) In: *Psicologia, laicidade e as relações com a religião e a espiritualidade*, v. 3: *Psicologia, Espiritualidade e Epistemologias Não Hegemônicas*. São Paulo: Conselho Regional de Psicologia de São Paulo – CRP 06, 2016a, p. 141-145.
- Marks, J., Copland, E., Loh, E., Sunstein, C. R., & Sharot, T. (2018). Epistemic spillovers: Learning others' political views reduces the ability to assess and use their expertise in nonpolitical domains. *Harvard Public Law working paper no. 18-22*, Harvard University, Cambridge, MA.
- Martin, C. & Nichols, R.: Personality and religion beliefs. *Jr. of Soc. Psychol.*, 1962, 56, 3–8
- Miles, J. & Shavlin, M. (2001). *Applying regression and correlation: a guide for student and researchers*, Sage, London
- Miot, H., A. (2017). Avaliação da normalidade dos dados em estudos clínicos e experimentais. *Jornal Vascular Brasileiro*, 16(2), 88-91. <https://dx.doi.org/10.1590/1677-5449.041117>

- Morier, D., & Keeports, D. (1994). Normal science and the paranormal: The effect of a scientific method course on students' beliefs. *Research in Higher Education*, 35, 443-453. 10.1007/BF02496382.
- Nechita, F., Alexandru, D.O, Turcu-Stiolica, R., & Nechita, D. (2015). Original Paper The Influence of Personality Factors and Stress on Academic Performance. *Current Health Sciences Journal*, 41, 47-61. 10.12865/CHSJ.41.01.07.
- Observatório do Plano Nacional de Educação (2020). *Indicadores e metas*.
<https://observatoriodopne.org.br/>
- Pais-Ribeiro, J., L. (2013). Medida na avaliação psicológica. *Psicologia, Saúde & Doenças*, 14(1), 245-263.
- Pasquali, L. (2007). Validade dos testes psicológicos: será possível reencontrar o caminho?. *Psicologia: Teoria e Pesquisa*, 23(spe), 99-107. <https://doi.org/10.1590/S0102-37722007000500019>
- Pervin, L. A., & John, O. P. (2004). *Personalidade: teoria e pesquisa*. Porto Alegre: Artmed
- Pino, F. A. (2014) A questão da não normalidade: uma revisão. *Revista de Economia Agrícola*, São Paulo, v. 61, n. 2. p. 17-33.
- Popper, K., (1972) *A Lógica da pesquisa científica*. São Paulo: Cultrix
- Rigdon, E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 3, 369-379. 10.1080/10705519609540052.
- Rutjens, B. T., Sutton, R. M., & van der Lee, R. (2018). Not all skepticism is equal: Exploring the ideological antecedents of science acceptance and rejection. *Personality and Social Psychology Bulletin*, 44(3), 384–405. <https://doi.org/10.1177/0146167217741314>
- Rutjens, B. T., van der Pligt, J., & van Harreveld, F. (2009). Things will get better: The anxiety-buffering qualities of progressive hope. *Personality and Social Psychology Bulletin*, 35, 535–543.

- Rutjens, B. T., van der Pligt, J., & van Harreveld, F. (2010a). Deus or Darwin: Randomness and belief in theories about the origin of life. *Journal of Experimental Social Psychology*, 46, 1078–1080.
- Rutjens, B. T., van Harreveld, F., & van der Pligt, J. (2010b). Yes we can: Belief in progress as compensatory control. *Social Psychological and Personality Science*, 1, 246–252
- Rynes, S., Colbert, A., & O'Boyle, E. (2018). When the “Best Available Evidence” Doesn't Win: How Doubts About Science and Scientists Threaten the Future of Evidence-Based Management. *Journal of Management*. 44. 014920631879693. 10.1177/0149206318796934.
- Sagan, C. (2006) *O mundo assombrado pelos demônios: a ciência vista como uma vela no escuro*. São Paulo: Companhia das Letras.
- Sanchez, C., Sundermeier, B., Gray, K., & Calin-Jageman, R. J. (2017). Direct replication of Gervais & Norenzayan (2012): No evidence that analytic thinking decreases religious belief. *PloS one*, 12(2), e0172636. <https://doi.org/10.1371/journal.pone.0172636>
- Schünemann, B. (2012). O juiz como um terceiro manipulado no processo penal? Uma confirmação empírica dos efeitos perseverança e correspondência comportamental. Trad. José Danilo Tavares Lobato. *Revista Liberdades*, n. 11, p. 30-50,
- Shenhav, A., & Greene, J. (2011). Divine Intuition: Cognitive Style Influences Belief in God. *Journal of experimental psychology. General*. 141. 423-8. 10.1037/a0025391.
- Sireci, S. G. (2005). Using bilinguals to evaluate the comparability of difference language versions of a test. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 117-138). Mahwah, NJ: Lawrence Erlbaum.
- Stagnaro, M. N., Ross, R. M., Pennycook, G., & Rand, D. G. (2019). Cross-cultural support for a link between analytic thinking and disbelief in god: evidence from India and the United Kingdom. *Judgment and Decision Making*, 14(2), 179-186.

- Ståhl, T., Zaal, M. P., & Skitka, L. J. (2016). Moralized rationality: Relying on logic and evidence in the formation and evaluation of belief can be seen as a moral issue. *PLoS One* 11:e0166332. doi: 10.1371/journal.pone.0166332
- Tabachnick, B.G. & Fidell, L.S. (2013) *Using Multivariate Statistics*. Pearson, Boston.
- Talhelm, T., Haidt, J., Oishi, S., Zhang, X., Miao, F. F., & Chen, S. (2015). Liberals Think More Analytically (More “WEIRD”) Than Conservatives. *Personality and Social Psychology Bulletin*, 41(2), 250–267. <https://doi.org/10.1177/0146167214563672>
- Valentini, F., & Damásio, B.F. (2016). Variância Média Extraída e Confiabilidade Composta: Indicadores de Precisão. *Psicologia: Teoria e Pesquisa*, 32(2), e322225. Epub October 27, 2016. <https://dx.doi.org/10.1590/0102-3772e322225>
- Vianna, H. (2015). Validade de construto em testes educacionais. *Estudos em Avaliação Educacional*. 24. 136. 10.18222/eae246020143309.
- Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior research methods*, 51(1), 409–428. <https://doi.org/10.3758/s13428-018-1055-2>

Apêndice B – Escala de Crença em Ciência

- 1 – A ciência nos fornece uma melhor compreensão do universo que a religião.
- 2 – Em um mundo assombrado por demônios, a ciência é uma vela no escuro (Carl Sagan).
- 3 – Só podemos acreditar racionalmente no que é cientificamente comprovável.
- 4 – A ciência nos diz tudo o que há para saber sobre a realidade.
- 5 – Todos os obstáculos que os humanos enfrentam podem ser resolvidos com a ciência.
- 6 – O método científico é o único caminho confiável para o conhecimento.
- 7 – O único tipo real de conhecimento que podemos ter é o conhecimento científico.
- 8 – A ciência é a parte mais valiosa da cultura humana.
- 9 – A ciência é o meio mais eficiente de alcançar a verdade.
- 10 – Os cientistas e a ciência devem ter mais respeito na sociedade moderna.

**Manuscrito 3 – Uma Decisão com Classe: Análise de Classes Latentes e
Indícios de Validade em Testes Situacionais de Julgamento**

**A Classy Decision: Latent Class Analysis and Evidence of Validity in Situational
Judgement Tests**

Thiago Augusto Costa de Olival
Universidade de Brasília

Resumo

Testes de julgamento e decisão são populares em pesquisas em áreas de intercessão entre a Economia e a Psicologia. Este manuscrito apresenta a construção de cenários de decisão envolvendo decisões financeiras e/ou em gestão de pessoas. Os cenários foram manipulados no que tange a qualidade e direção das informações, de modo que a amostra de gestores (n = 268) foi dividida de maneira aleatória para que fossem comparados os padrões de resposta entre os grupos. Os indícios de validade são observados por meio da discriminação dos indivíduos pela utilização da técnica da Análise das Classes Latentes, além da comparação entre os grupos e a correlação entre os padrões de resposta observados nos cenários (decisões tidas como corretas) com as variáveis hipotetizadas neste trabalho (crença na ciência, a racionalidade, o tempo de experiência na empresa, a idade e traços de personalidade, além das influências envolvendo a quantidade e qualidade das informações).

Palavras-chave: Julgamento e decisão organizacional, Análise de classes latentes, testes situacionais de julgamento e decisão

Abstract

Judgment and decision-making tests are popular in research in areas of intersection between Economics and Psychology. This manuscript presents the construction of decision scenarios involving financial decisions and / or people management. The scenarios were manipulated to the quality and direction of the information, so that the sample of managers (n = 268) was randomly divided so that the response patterns between groups were compared. The signs of validity are observed through the discrimination of individuals through the use of the Latent Class Analysis technique, in addition to the comparison between groups and the correlation between the response patterns observed in the scenarios (decisions considered to be correct) with the variables hypothesized in this study (belief in science, rationality, experience in the company, age and personality traits, in addition to influences involving the quantity and quality of information).

Keywords: Judgement and decision-making, Latent Class Analysis, Situational Judgement test

Introdução

Dentre os muitos objetivos organizacionais relacionados a gestão de pessoas, considera-se que a classificação dos indivíduos de acordo com suas características é de extrema importância, visto que pode ser primordial para processos envolvendo seleção de pessoal, avaliação de lacunas de desenvolvimento além de ordenar e qualificar os indivíduos em relação a performance no trabalho. Por meio da classificação é possível discriminar as características dos indivíduos e relacioná-las com algum indicador, observando, dentre outros aspectos, a probabilidade de que um indivíduo apresente características ou padrões observados em alguma classe.

A qualificação de gestores já se mostrou um processo complexo e que muitas vezes tem cunho subjetivo no que tange a avaliação em relação a performance e resultados de suas ações (Cohen, 2019). Desta forma, testes que possam reduzir a subjetividade na identificação de perfis relacionados a algum tipo de competência organizacional ou padrão de desempenho em situações aderentes as realidades do mundo do trabalho podem auxiliar as organizações a decidir sobre capacitar ou contratar indivíduos com maior precisão. Estudos desta natureza ainda são escassos visto que testes situacionais mais estruturados e com indícios de validade no Brasil ainda não apresentam grande popularidade.

O presente estudo tem por objetivo adaptar, criar e identificar indícios de validade de um teste situacional envolvendo cenários de decisão. Os cenários apresentados no teste descrevem situações relacionadas ao universo das finanças e ao universo da gestão de pessoas em organizações. As situações-problema desenvolvidas para estímulo ao processo decisório em cenários de finanças envolvem avaliação sobre aplicação ou a retirada de recursos, escolhas de investimentos sob a ótica probabilística e majoração de resultados. Para os cenários envolvendo gestão de pessoas, são utilizados conceitos relacionados a aplicação e julgamento disciplinar decorrente de falha funcional, avaliação, seleção e contratação de candidatos considerando opções de perfis e relatos de comportamento dos candidatos. Para alcançar o objetivo exposto, o presente manuscrito descreve procedimentos de análise realizados por meio da análise das classes latentes e por meio

das relações entre os padrões de resposta e as variáveis presentes no estudo (traços de personalidade, crença em ciência, racionalidade, experiência profissional e idade). As classes latentes são análogas a análise fatorial confirmatória, de maneira que auxiliam a caracterizar indícios de evidência da qualidade da estrutura interna do instrumento.

Referencial Teórico

A Utilização de Testes Situacionais de Julgamento

Testes situacionais de julgamento apresentavam popularidade nos Estados Unidos quando comparados com outros países na primeira década dos anos 2000. Tais questionários eram constantemente utilizados em processos de seleção de pessoal (Ployhart, 2006) pois permitem medir habilidades em grandes grupos ainda em estágio inicial do processo seletivo.

Os testes permitem emular situações reais de trabalho, sendo apresentados de maneira escrita ou com a utilização de multimídia e, como resultado, é esperado que o indivíduo faça escolhas ou ordene itens de acordo com as orientações do instrumento, simulando muitas vezes um processo de entrevista (Lievens et al, 2008) ou uma deliberação de algum comitê. Os testes podem ser de avaliação de conhecimento ou comportamento, muitas vezes se assemelhando a testes de inventários de personalidade (McDaniel et al, 2007) ou se ramificando (Olson-Buchanan et al, 1998), em que os indivíduos se confrontam com consequências ou desdobramentos de suas escolhas.

Uma das premissas observadas sobre testes situacionais é que auxiliariam a entender as relações entre variáveis e classificação dos indivíduos em relação a sua aptidão para determinada competência. Importante ressaltar que alguns dos testes mencionados possuem elevada complexidade e trazem a possibilidade de observação de múltiplos domínios (Lievens et al, 2008).

Com o intuito de entender e classificar as respostas entre certo ou errado, algumas das técnicas utilizadas para este fim pressupõem a aplicação de testes em grupos de funcionários considerados especialistas, em que as melhores respostas seriam pontuadas como corretas e as piores como incorretas (Bergman et al, 2006; Lievens et al, 2008).

Alguns dos benefícios percebidos como resultados do procedimento de criação e aplicação de testes situacionais elencados na literatura estão relacionados aos custos de desenvolvimento em relação a outras práticas de seleção similares (Lievens et al,2008), predição de desempenho no trabalho (Motowildo et al,1990) e um melhor conhecimento das variáveis envolvidas com a situação-problema e as relações destas com a medida que se deseja observar (Whetzel & Mc Daniel, 2009). Desta forma, testes situacionais são uma opção factível para mensurar a combinação de variáveis como personalidade, experiências, crenças, conhecimentos acadêmicos e relacionados ao trabalho tendo como base um determinado cenário de decisão (Motowildo & Beimer, 2010; Patterson et al, 2019).

A utilização dos testes na área de saúde aparentemente é mais comum, pela quantidade de artigos publicados em revistas nas áreas da medicina, farmácia e enfermagem que tratam do tema em contexto de seleção de profissionais, visto a quantidade de manuscritos encontrados em plataformas de consulta deste tipo de manuscrito. Tais estudos consideram a relação do fenômeno estudado com medidas e construtos diversos, mas não se restringem somente a tal cenário visto que há artigos discutindo e utilizando os testes em áreas como a educação e literacia financeira (Fernandes et al, 2014; Kaiser & Menkhoff, 2017; Wuttke et al, 2020).

Testes Situacionais Envolvendo Decisões Financeiras

Quando se fala em julgamento e decisão aplicada ao contexto de finanças, economia comportamental ou psicologia econômica, algumas das medidas observadas sobre decisão em contexto de investimento derivam das propostas desenvolvidas por Kahneman e Tversky entre as décadas de 70 e 80. No Brasil, dentre algumas das versões adaptadas, muitas são semelhantes as observadas nos estudos tidos como exemplo de Kimura et.al, (2006), Silva et al. (2009) e Soares e Barbedo, (2013). Costumeiramente tais testes apresentam cenários com enunciados curtos com decisões muitas vezes genéricas e pouco aplicadas a um contexto decisório organizacional real, em que fatores como a quantidade e qualidade das informações poderia exercer alguma influência. Como resultado das análises, muitas vezes

são propostas as comparações das médias ou frequências entre os grupos com base no pressuposto teórico da literatura específica relacionada a julgamento e decisão.

Bases Para a Decisão

Decisões racionais são requeridas em todas as organizações e Kang e Park (2019) relembram que decisões no sistema financeiro tem altos impactos econômicos para a sociedade. Ressaltam também aos pesquisadores que o setor e o ambiente de negócios dos bancos mudou consideravelmente nos últimos anos e que conforme a complexidade do cenário se eleva, mais desafiador se torna o processo de tomada de decisão.

A tomada de decisão nas organizações é processo comum em cargos gerenciais, de forma que muito deste processo é realizado de maneira informal (Cohen 2019) e nem sempre são utilizadas as evidências disponíveis (Tort Martorell et al, 2011). Muitas vezes, as informações disponíveis também podem ser ignoradas ou mal interpretadas (Rynes et al 2018), de maneira que fatores externos (estratégias empresariais ou da equipe, cenários econômicos, ambiente político, etc.) exercem influência relevante para no contexto decisório (Katapol, 2018).

Dentre as muitas variáveis relacionadas ao indivíduo e possíveis de compor um modelo envolvido em uma decisão mais racional, há indícios de que traços de personalidade (Hitt & Tyler, 1991; Mcelroy & Dowd, 2007; Furnham & Boo, 2011; Caputo, 2014; Teovanovic, 2019), racionalidade (Antonczyk & Salzman, 2014; Cheel & Noren, 2017; Katopol, 2018) e experiência profissional (Yustina & Gudono, 2017) possam influenciar o processo decisório. Ainda, a relação entre a performance de homens e mulheres em testes envolvendo finanças e literacia financeira já se mostrou significativa. Wuttke et al., (2020) discutem acerca das diferenças relacionadas a este fator e observadas de maneiras distintas, dependendo do país onde o teste fora aplicado.

Desta forma, a combinação de múltiplas variáveis como as expostas vão ditar os resultados da decisão em ambiente de racionalidade limitada (Simon, 1955), mas os caminhos possíveis são muitos. Embora os fatores externos sejam os mesmos para agrupamentos semelhantes de gestores, como por exemplo em uma situação em que todos

indivíduos estejam em uma mesma unidade organizacional, cada indivíduo vai interpretar o cenário de forma peculiar, de acordo com variáveis comuns a todos (como personalidade) ou mais específicas, como conhecimento técnico e experiência profissional (Yustina & Gudono, 2017).

Os vieses que podem influenciar a decisão são muitos. Para o presente estudo, são elencados alguns que serão alvo de avaliação, como o efeito Halo no julgamento envolvendo pessoas (Bol & Smith, 2011; Yustina & Gudono, 2017), tendências culturais ou decorrentes do conhecimento sobre determinado tema (Czerwonka, 2017), o excesso de confiança (Antonczyk & Salzmann, 2014), maior susceptibilidade a efeitos heurísticos decorrentes de forma de pensar menos analítica (Tversky & Kahneman, 1974; Cheek & Norem, 2017) e até mesmo a quantidade e qualidade das informações presentes no cenário (Yang et al., 2018)

Indícios de Validade em Testes Situacionais

Muitos testes de julgamento são elaborados de maneira complexa em virtude de sua multidimensionalidade e tal característica pode impactar na observação do alfa de Cronbach e na extração de fatores – métodos tradicionais para aferir confiabilidade e validade (Schmitt & Chan, 2006; Schreurs et al, 2020). Alguns dos indícios de validade relacionadas a critério e validade incremental pressupõem relações de predição com desempenho do trabalho, traços de personalidade, capacidade cognitiva e inteligência prática (McDaniel et al 2007; Lievens et al, 2007). A discussão acerca dos indícios de validade também pode perpassar sobre a manipulação do teste a fim de que se observe que a variável latente muda conforme o teste é alterado (Baghaei & Yazdi, 2016).

Não há evidências que indiquem que acertos em testes de decisão mais genéricos estejam relacionados a acertos em testes mais complexos, e embora algumas das variáveis sejam comuns a ambos, há outras habilidades requeridas conforme a complexidade e aderência a realidade dos testes aumenta. Dentre alguns dos exemplos observados na literatura, Kang e Park (2019) propuseram cenários aplicados ao contexto de decisões que são aderentes a realidade de processos presentes no setor bancário para testar efeitos

heurísticos. Embora tenham sido observadas diferenças entre médias na comparação entre amostras e cenários, tais cenários ou itens não apresentam indícios de validade tradicionais.

Desta forma, as discussões acerca da validade dos testes situacionais são apresentadas por Wuttke et al, (2020) e são exatamente sobre o fato de que as diferenças entre grupos de comparação estejam atreladas as baixas qualidades psicométricas do instrumento e não as medidas utilizadas por cada indivíduo com o intuito de realizar a escolha correta. Alguns dos indícios de validade sugeridos para testes desta natureza perpassam pela confiabilidade, validade de critério, validade incremental, validade de construto, percepções sobre a aplicação, susceptibilidade ao treino/prática, falseabilidade e utilidade (Lievens et al., 2008; Schreurs et al, 2020). Embora estes não sejam os únicos indícios, são alguns dos mais observados e não necessariamente obtidos de maneira uniforme (utilizando as mesmas técnicas em todos os instrumentos).

Para que sejam avaliados indícios de validade de conteúdo, o instrumento deve ser submetido a juízes para que observassem as qualidades no que tange a relevância, clareza e a pertinência dos itens. Alexandre e Coluci (2011) sugerem que é desejável um índice de convergência superior aos 90% entre os juízes para que exista indicio de validade de conteúdo. Outra possibilidade para identificar indícios de validade em testes situacionais é a análise de classes latentes que permite a discriminação dos indivíduos. Uma das possibilidades de investigar a heterogeneidade populacional (observando a probabilidade de um indivíduo pertencer a um grupo ou classe) além da relação das variáveis presentes no estudo é a realização da análise de classes latentes (McLachlan & Peel, 2000; Borbsoom, 2008; Collins & Lanza, 2010). A vantagem deste tipo de análise é que pode comportar variáveis contínuas e categóricas de maneira simultânea (Maysn et al, 2010, Muthén & Muthén, 2014), de maneira que não se trata necessariamente de um grupo pontuar mais do que outro em determinados itens, mas sim de sugerir um perfil qualitativo (ou um padrão), discriminando os indivíduos. (Hauck, 2015).

As meta-análises nesta área ainda são escassas e Webster et al, (2020) apresenta uma discussão a cerca das observadas evidências de validade em múltiplos estudos

geralmente na área de saúde, no que diz respeito ao uso de variáveis não acadêmicas como preditores dos resultados do teste. Mesmo considerando um montante de artigos inicialmente observado pelos pesquisadores na casa dos 300 artigos, o valor final de manuscritos utilizados para fins da meta-análise não passou dos 20 artigos, o que sugere que apesar da possibilidade de evidências de que testes desta natureza possuam indícios de validade, considera-se que exista um longo caminho até que tais evidências sejam mais robustas ou que ao menos existam maiores possibilidades de comparar testes desta natureza.

O presente estudo busca identificar possíveis indícios de validade dos testes situacionais desenvolvidos para aplicação em uma amostra de gestores, decorrentes da discriminação destes indivíduos em classes agrupadas por traço latente, além de observar as relações dos padrões de resposta apresentados nos cenários com variáveis externas (personalidade, crença, racionalidade, sexo e experiência profissional).

Método

Construção do Teste de Decisão Organizacional Aplicado a Finanças e a Gestão de Pessoas

O objetivo da construção do presente instrumento é observar o comportamento e eventuais padrões de decisão dos indivíduos quando apresentados a múltiplos cenários que simulam possibilidades de escolha comuns ao contexto organizacional em decisões envolvendo finanças ou gestão de pessoas. Além disso, o instrumento tem o propósito de testar se existe a possibilidade de classificar os participantes de acordo com padrões de decisão ou raciocínio que os diferenciem, de modo que os resultados podem apresentar a probabilidade de que o indivíduo pertença a determinada classe. Tais padrões de pensamento e decisão seriam manifestados nos testes por meio de traços latentes. Como forma de avaliar possíveis alterações na observação dos padrões de decisão, alguns dos testes foram objeto de manipulação em relação a induzir a determinados padrões de resposta ou suscitar efeitos heurísticos (algumas informações foram modificadas de acordo

com o grupo, afim de comparar se existiram diferenças significativas entre o grupo considerado como controle e o grupo de teste).

Descrição dos Cenários

O teste proposto apresenta dois tipos de cenários para tomada de decisão: finanças e gestão de pessoas. O motivo da escolha do conteúdo dos testes diz respeito a hipótese de que o pensamento mais racional seria mais facilmente manifestado pelo conteúdo que aborde decisões financeiras. Os cenários foram assim separados: FINAN1A, FINAN1B, FINAN2, FINAN3; PESS1A, PESS1B, PESS1C, PESS2 e PESS3, sendo 3 cenários gerais para cada conteúdo proposto. Os cenários foram apresentados aleatoriamente aos participantes da pesquisa, o que os separa por grupos de respondentes: Grupo1, Grupo2, etc.

As questões FINAN1A e FINAN1B tem como objetivo observar a capacidade de avaliação probabilística dos sujeitos mediante a escolha de eventos mais prováveis de ocorrer, considerando o objetivo de minimizar o risco. Neste grupamento de questões, os indivíduos são confrontados com o caso da empresa ABC e se deparam com duas decisões. Na primeira pergunta, é apresentada uma medida de probabilidade de ganho, enquanto que na segunda há uma probabilidade de perda em relação ao ganho. Para o grupo 1, foi disponibilizado o cenário em que há sugestão de que as decisões anteriores da empresa são boas (há um prospecto anterior positivo) enquanto que no segundo cenário, há um prospecto ruim. O grupo 2 recebeu as questões na mesma ordem, mas com uma alteração no que tange ao prospecto anterior dos investimentos: na primeira questão o prospecto anterior é ruim e na segunda é bom, como mecanismo de comparar se haveria influencia tanto da ordem quando da sugestão de eventos anteriores no padrão de resposta.

Na questão FINAN2, os indivíduos recebem o caso de uma construtora que no passado teve problemas com o crime de lavagem de dinheiro. A citação do evento negativo (envolvimento de dois sócios e investigação de um terceiro) é a maneira de induzir que a empresa, possa conter outros indícios de tal prática. O crime é conceituado pela lei 9613/1998 - remodelada pela 12683/2012 - que trata de forma complexa os crimes contra o

sistema financeiro que versam sobre injetar na economia recursos de origem ilícita oriundos de sonegação de impostos ou de práticas criminosas. Algumas das muitas formas que tal delito ocorre envolvem a compra de bens e dissimulação ou ocultação da origem do recurso de maneira a fazer com que os recursos se tornem lícitos. Ressalta-se que algumas atividades econômicas são consideradas como mais sensíveis e demandam dos envolvidos (fiscalizadores, reguladores, bancos e outros interessados) maior controle das atividades e observância das atividades financeiras. Tal pressuposto está atrelado às características da atividade econômica que podem dificultar a rastreabilidade da origem e do destino do recurso (como igrejas), o comércio de bens de valor elevado ou luxo (loja de carros) ou fraudes imobiliárias (casas, apartamentos, etc.). Nesta questão, a indicação correta é a escolha que apresenta a maior projeção de lucros, visto que não há indícios de que o investimento tenha ilicitudes e, obviamente a prática econômica de majoração de lucros não é criminosa. Desta forma, o enunciado da questão busca dificultar uma escolha e trazer elementos que induzem a uma suposição de ilícito, mas que não se relacionam com o objetivo da decisão.

A questão FINAN3 busca observar o apego emocional do indivíduo em relação a alocação de recursos em projetos, considerando que o projeto alvo da decisão é relatado como sendo um “sonho pessoal e compartilhado com a alta administração”. Porém, o retorno está abaixo do esperado e é necessário realizar uma decisão sobre reinvestir no projeto ou abandonar e partir para outro projeto. No primeiro cenário, o apelo é sobre o resultado do projeto que está muito abaixo do esperado e os recursos foram consumidos quase que na integralidade, e no segundo o resultado ainda permanece aquém, mas não de maneira tão extrema. Sob a ótica da teoria do custo afundado, é esperado que os indivíduos de forma racional deixem de lado o apego emocional e partam para um novo projeto, em especial em cenários mais críticos em que as perdas já são elevadas ou iminentes.

O grupamento que envolve as questões PESS1A, PESS1B e PESS1C apresentam a situação de dois funcionários que cometem falhas operacionais que ocasionam prejuízos financeiros. Em PESS1A, na apresentação do caso, valores de prejuízo são menores para o

grupo 1 e maiores para o grupo 2 com objetivo de avaliar se os valores envolvidos são de fato relevantes para decisões disciplinares ou se o simples fato consumado ou o histórico já seriam suficientes para um desfecho. A decisão envolve avaliar se a demissão da personagem foi correta. Em PESS1B, apresenta-se o caso de outra personagem que, apesar do histórico de boa conduta, foi responsável por um prejuízo igual (para o grupo 1) ou maior (para o grupo 2) quando comparado ao caso da personagem inicialmente citada em PESS1A. Neste cenário, é necessário ponderar se a demissão foi acertada e há possibilidade da decisão ser ancorada na decisão do cenário PESS1A. Por fim, na questão PESS1C, o indivíduo deve escolher qual dos dois indivíduos foi corretamente demitido e vai ponderar as decisões feitas em PESS1A e PESS1B sobre a demissão daquele que julga como merecedor do desfecho.

A questão PESS2 trata da seleção de um gestor para desempenhar funções que sejam demandadas habilidades de liderança, sendo esta característica transcrita tanto no texto quanto na descrição de um dos candidatos. Porém, as três opções possuem nomes da história mundial com características positivas e negativas. A sugestão dos nomes pode induzir o indivíduo a deixar de lado o objetivo de selecionar alguém que tenha a característica “habilidades de liderança” e apegar-se aos nomes e representação que as figuras históricas possuem, de maneira a gerar um preconceito. Obviamente, buscou-se apresentar figuras históricas que remetam a liderança, sendo duas com um apelo mais negativo (Nero e Adolf Hitler) e uma positiva (Martin Luther King). Embora os nomes tenham esta proposta, não se tratam das figuras em si e a escolha correta seria o candidato Adolf.

Na questão PESS3, há um dilema envolvendo uma escolha mais relacionada a opções conservadoras ou arrojadas. Não se trata de uma decisão certa ou errada propriamente dita, mas uma avaliação quanto a conservadorismo (relacionada a aversão a risco do ente racional e a escolha de alguém já experiente) ou ter um perfil mais arrojado em decisões envolvendo pessoas e escolher um novo talento. O dilema trata sobre promover um funcionário e as opções são sobre alguém que já foi gestor e tem alguma referência ou uma aposta de um novo talento.

A tabela 1 resume as principais inserções e as diferenças entre os cenários, além de apresentar os respectivos objetivos de cada questão. O que se demonstra é que em algumas questões, o padrão de informação se altera no enunciado e pode induzir a algum efeito heurístico e/ou suscitar avaliações probabilísticas ou efeitos de apego emocional durante o processo de decisão.

Tabela 1

Resumo das questões e induções inseridas nos cenários

Cenário	Grupo 1	Grupo 2	Relação	Objetivo
FINAN1A	Bom	Ruim	Finan1a e Finan1b	Avaliação probabilística de ganhos
FINAN1B	Ruim	Bom	Finan1a e Finan1b	Avaliação probabilística de ganhos com possibilidade de perda
FINAN2	Igual	Igual	Independente	Escolha entre projetos; heurística da representatividade
FINAN3	Arrojado	Moderado	Independente	Decisão referente a investimento com apelo emocional e aquém de retorno esperado (custo afundado)
PESS1A	\$10.000	\$100.000	Pess1a, Pess1b e Pess1c	Avaliação sobre a demissão de dois funcionários. Heurística de ancoragem em relação ao valor além do prospecto
PESS1B	Igual	Igual	Pess1a, Pess1b e Pess1c	Comparação com base nos valores do cenário anterior
PESS1C	Igual	Igual	Pess1a, Pess1b e Pess1c	Decisão final em relação ao caso
PESS2	Igual	Igual	Independente	Escolha entre 3 candidatos. Indução a preconceito (Halo) e Heurística da representatividade
PESS3	Igual	Igual	Independente	Decisão entre contratação de dois perfis. Arrojado (candidato mais novo) ou conservador (Candidato mais antigo).

Participantes

Participaram da presente pesquisa 268 gestores de uma instituição financeira, sendo que os respondentes foram divididos aleatoriamente em dois grupos de 136 e 132 indivíduos. A amostra é majoritariamente constituída por participantes do sexo masculino (68%), com escolaridade predominante de ensino de pós-graduação (89%), com intervalo de idade predominante variando de 41 a 50 anos. A distribuição geral da amostra é apresentada conforme abaixo, na tabela 2:

Tabela 2*Caracterização dos participantes da pesquisa*

Características	Distribuição		
	<i>n</i>	%	
Sexo	Masculino	183	68%
	Feminino	85	32%
Escolaridade	Nível Médio	4	1%
	Superior-Graduação	10	6%
	Superior -Pós-Graduação - Especialização	246	89%
	Superior – Mestrado	6	3%
	Superior - Doutorado	2	1%
Idade	31-35	17	6%
	36-40	49	18%
	41-45	73	27%
	46-50	54	20%
	51-55	57	21%
	56-60	17	6%
	61-65	1	0%
Tempo de empresa	Até 10 anos	8	3%
	10-15	52	19%
	16-20	105	39%
	21-25	24	9%
	26-30	23	9%
	30+	56	21%

Procedimentos de Coleta de Dados

A pesquisa foi realizada no ano de 2020 em uma organização do sistema financeiro com abrangência nacional. A coleta foi realizada pelo aplicativo *Microsoft Forms* cujo link foi enviado por *e-mail* aos possíveis respondentes. Na instituição, existem basicamente dois níveis principais de gestores: a gerência média, que são gestores em um primeiro ou segundo nível hierárquico e os administradores, que coordenam múltiplas equipes de maneira a representar os interesses da instituição em suas áreas de atuação, sendo superiores hierárquicos da gerência média. Para o estudo, serão selecionados os administradores das unidades envolvidas. Os participantes foram convidados a participar da pesquisa de maneira aleatória, considerando que a amostra foi dividida em dois grupos de aproximadamente 150 indivíduos.

Procedimentos de Análise de Dados

Para que fossem avaliados indícios de validade de conteúdo, o instrumento foi submetido a juízes para que fossem observadas as qualidades no que tange a relevância, clareza e pertinência dos itens. Em uma escala de 1 a 4, 6 juízes (membros de um grupo de

pesquisa composto por mestrandos e doutorandos com formação em administração, economia e psicologia). O pressuposto em questão estaria de acordo com a sugestão de Alexandre e Coluci (2011) de maneira que é desejável um índice de convergência superior aos 90% entre os juízes para que exista indício de validade de conteúdo, considerando a quantidade de notas 3 e 4 em relação a quantidade de itens. O grupo de juízes teve índice apurado em 97%.

Para análise dos padrões de respostas no teste situacional, foi realizada a análise de classes latentes. Dentre os parâmetros de ajuste do modelo se observam o *Akaike Information Criterion* – AIC, o *Bayesian Information Criterion* – BIC e a estimação realizada pelo método da máxima verossimilhança Maximum Likelihood, de maneira que o melhor modelo é aquele que apresenta os menores valores na comparação. Os resultados dizem respeito a probabilidade de que uma classe identificada apresente determinada pontuação ou padrão em relação a um item. Também é necessário avaliar a significância do teste que compara os modelos, de forma que a hipótese nula é que o modelo com menos classes tem melhor ajuste que o modelo com mais classes. Por fim, um indicador importante na comparação é o coeficiente de entropia, que varia de 0 a 1 e indica a qualidade da classificação, sendo desejáveis valores próximos de 1 (Nylund et al, 2007; Collins & Lanza, 2010).

Por fim, para avaliar os indícios de validade de critério, foi realizada uma comparação entre os padrões de respostas específicos dos grupos de respondentes relacionando os itens do questionário (no caso, a resposta considerada correta ou racional) com algumas das variáveis sugeridas na literatura e citadas neste manuscrito onde já foram observadas relações significantes (crença em ciência, traços de personalidade, capacidade cognitiva e inteligência).

Resultados

Análise das Classes Latentes

Para a Análise das Classes Latentes, foram observados os parâmetros de comparação entre modelos nos indicadores do BIC, *Loglikelihood*, entropia e a possibilidade de interpretação com base na teoria a qual foi baseada (Nylund et al, 2007; Collins & Lanza, 2010). Desta forma, conforme se observa na tabela 3, apresentam-se os modelos considerando 2, 3 e 4 classes, dos quais optou-se pelo modelo 1(2 classes) sob a ótica da interpretação teórica e estatística.

Tabela 3

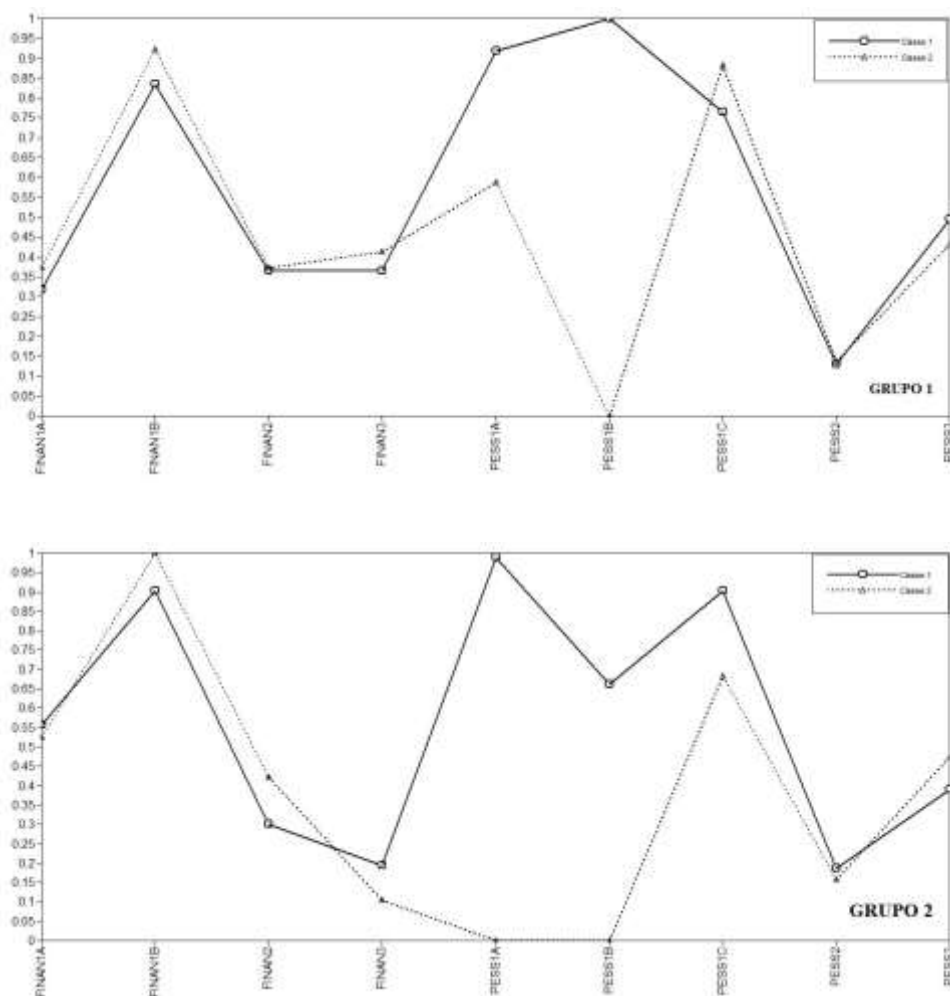
Comparação entre os modelos da pesquisa

	<i>Modelo</i>							
	Classe 1	Classe2	Classe3	Classe4	LL	BIC	Entropia	X ²
Grupo 1								
Modelo 1	85(62%)	51(38%)	-	-	-785,4	1683,8	1	792,12
Modelo 2	40(29%)	50(37%)	46(34%)	-	-769,1	1710,2	0,96	680,15
Modelo 3	39(29%)	38(28%)	15(11%)	44(32%)	-756,1	1743,1	0,95	629,35
Grupo 2								
Modelo 1	113(86%)	19(14%)	-	-	-703,3	1518,9	0,98	588,99
Modelo 2	102(77%)	11(8%)	19(14%)	-	-689,6	1550,1	1	555,36
Modelo 3	32(24%)	22(17%)	59(45%)	19(14%)	-679,6	1588,7	0,92	362,72

A separação em classes possibilitou identificar padrões de respostas distintos entre as classes, de maneira que as principais diferenças se apresentaram nas questões relacionadas a gestão de pessoas. A figura 1 apresenta os padrões das classes em ambos os grupos. Os padrões de respostas mais racionais (acertos) ou intuitivas (erradas) se apresentam de maneira semelhante entre as classes e entre grupos (pessoas que receberam o cenário em uma direção de informação em relação as pessoas que receberam os cenários em outra direção de informação), mas as diferenças se evidenciam mais nas questões envolvendo pessoas (PESS1A, PESS1B e PESS1C) indicando que gestores pertencentes as classes 1 e 2 comportam-se de maneira distinta quando o julgamento está relacionado a cenários envolvendo pessoas e prejuízos financeiros.

Figura 1

Comparativo da distribuição das classes entre grupos



Observa-se também que além da discriminação em classes, há alteração do padrão das respostas nas classes nos cenários em que há alguma manipulação no que tange a qualidade das informações (FINAN1A, FINAN1B, FINAN3 e PESS1A) ou naqueles em que se trata de um desdobramento de um cenário principal (PESS1B e PESS1C em relação a PESS1A).

A classificação destes gestores foi batizada como Gestores orientados a finanças (Classe 2) e Gestores orientados a pessoas (Classe 1). Uma hipótese é que gestores com orientação a finanças tiveram resultados inferiores pois possivelmente ficaram presos ao custo afundado ou ancorados com o evento da demissão, de maneira que podem ter optado pela não demissão na tentativa de reaver de alguma forma os valores, enquanto que

gestores mais voltados a pessoas foram mais diretos e em uma tentativa de evitar eventos desta natureza no futuro, foram mais pontuais e assertivos na demissão imediata.

Comparação Entre Grupos e Classes

Os grupos foram separados aleatoriamente considerando as manipulações (por meio de informações inseridas) nos cenários dos testes. A comparação realizada tem como objetivo avaliar eventuais diferenças observáveis entre as classes e se as manipulações propostas podem diferenciar os grupos.

Por meio de um teste não paramétrico de duas amostras independentes (U de Mann-Withney) foram realizadas as comparações entre os grupos e as classes.

Para a comparação dos grupos (entre a amostra 1 e a amostra 2), observou-se que os cenários FINAN1A ($U=7048$; $p<0,05$) e FINAN3 ($U=7176$; $p<0,05$) apresentaram significância, indicando a diferenciação entre os grupos.

Já para as classes (classe 1 em relação a classe 2), a diferença na comparação foi observada em PESS1A ($U=3250$; $p<0,05$) e PESS1B ($U=1330$; $p<0,05$).

Correlação entre Padrão de Respostas e Variáveis

Com o objetivo de avaliar validade de critério, foram realizadas correlações entre as respostas aos cenários e variáveis indicadas pela literatura como passíveis de influenciar o julgamento. Desta forma, objetiva-se avaliar indícios de validade considerando a relação dos testes com variáveis externas e outros construtos (Crombach & Meekl, 1955; Campbel & Fiske, 1959; Brown, 1970; Pasquali, 2007; Pais-Ribeiro, 2013; Viana, 2015). Ao realizar uma correlação de Spearman utilizando Bootstrap com 1000 interações para um intervalo de confiança de 95%, foram identificadas algumas relações entre a crença na ciência, traços de personalidade, racionalidade, sexo, tempo de empresa e idade e o padrão de resposta ativado pelos cenários. As relações observadas estão descritas na tabela 4 e dizem respeito possível relação entre as variáveis elencadas no estudo e o padrão de respostas considerado como correto.

Tabela 4

Correlações entre cenários e variáveis do modelo

	CENÁRIO DO GRUPO 1								
	FINAN1A	FINAN1B	FINAN2	FINAN3	PESS1A	PESS1B	PESS1C	PESS2	PESS3
Crença em ciência	-0,079	-0,134	0,067	0,163***	0,075	-0,004	-0,046	0,127	-0,115
Abertura a Exp.	-0,119	0,008	0,089	-0,056	-0,081	-0,002	-0,071	0,109	-0,120
Amabilidade	0,093	0,235**	-0,113	-0,119	0,009	-0,061	-0,042	0,007	-0,115
Conscienciosidade	0,051	0,157***	0,024	-0,017	0,014	-0,061	0,043	-0,037	-0,155***
Extroversão	-0,108	0,035	0,092	0,003	0,002	-0,043	-0,057	0,045	-0,167***
Neuroticismo	0,066	-0,197	-0,005	0,028	-0,052	-0,064	0,053	0,060	0,077
Racionalidade	-0,024	0,113	-0,053	0,117	0,073	0,072	0,053	-0,097	0,023
Sexo	0,099	0,185	-0,077	-0,061	0,017	-0,155	0,080	0,012	-0,132
Tempo de empresa	0,055	-0,117	-0,011	0,222	0,003	-0,018	-0,036	0,034	0,086
Idade	0,123	-0,067	-0,120	0,159***	0,140	0,037	0,016	-0,007	0,026

	CENÁRIO DO GRUPO 2								
	FINAN1A	FINAN1B	FINAN2	FINAN3	PESS1A	PESS1B	PESS1C	PESS2	PESS3
Crença em ciência	0,142	-0,073	-0,092	0,103	-0,129	-0,037	-0,050	0,032	0,096
Abertura a Exp.	0,130	-0,130	-0,119	-0,032	0,064	0,058	-0,106	0,001	0,031
Amabilidade	-0,074	-0,141	-0,090	0,123	-0,169***	-0,201	0,001	-0,075	0,014
Conscienciosidade	0,116	-0,185	-0,062	0,037	-0,056	0,052	-0,068	-0,020	0,023
Extroversão	0,094	-0,116	-0,123	-0,057	0,062	0,014	-0,128	-0,101	0,063
Neuroticismo	0,041	0,146	-0,024	-0,098	0,146	0,185	0,100	0,154	-0,124
Racionalidade	-0,108	0,172	-0,079	0,146	-0,109	-0,005	-0,099	-0,026	-0,018
Sexo	0,155	-0,052	0,047	-0,037	-0,156	-0,058	0,061	-0,078	-0,086
Tempo de empresa	-0,007	-0,018	-0,022	-0,010	-0,129	-0,139	-0,054	-0,048	0,122
Idade	-0,069	-0,096	-0,119	-0,084	-0,173	-0,201	-0,073	-0,092	-0,151

Nota: ** Significante em $p < 0,01$; ***Marginalmente significativa (próximo ao limite para $p > 0,05$)

Os resultados observados na tabela 4 evidenciam que o padrão de resposta racional se relacionou com as variáveis e construtos presentes no modelo. As relações foram observadas individualmente, para comparação com as eventuais manipulações realizadas. Dentre as principais relações, a personalidade e seus fatores foram os que apresentaram a maior quantidade de observações (amabilidade = 3; conscienciosidade = 3; neuroticismo = 2; extroversão = 1 e abertura = 0), seguidos do sexo com 4 observações, crença em ciência com 1 e idade com 2 observações, além de racionalidade e tempo de empresa com uma observação cada.

Discussão

Os resultados do estudo corroboram que os testes situacionais se mostraram uma forma de avaliar e classificar os indivíduos de maneira peculiar e elevam as possibilidades de discriminar de maneira ímpar as características e padrões de respostas. Os cenários apresentam evidências de que é possível classificar gestores de acordo com padrões de decisão, de maneira que a discriminação de classes mais sensível ocorreu em cenários em que as decisões envolviam conjuntamente pessoas e valores, de forma que alguns gestores

podem ter sofrido consequências de ancoragem na expectativa de recuperar os valores enquanto que outros decidiram pelo encerramento do problema de maneira mais tempestiva. O coeficiente de entropia apresentado sugere que a classificação observada é muito boa em ambos grupos, visto que seu índice é igual ou muito próximo a 1, indicando excelente classificação (Ramaswamy et al, 1993; Celeux & Soromenho, 1996; Collins e Lanza 2010). Sugerimos que a classificação de maneira robusta corroborada pelo índice de entropia evidencia um indício de validade de critério (predição do indivíduo ao pertencimento de classes e seu consequente padrão de respostas mais intuitiva ou mais racional e com melhor performance em cenários envolvendo finanças ou em cenários envolvendo pessoas).

Quanto aos indícios de validade de critério, foram identificadas correlações importantes entre as respostas e variáveis externas. Algumas das relações observadas entre variáveis individuais (personalidade, experiência, crenças, etc), externas (quantidade e qualidade de informações disponíveis) e o objeto de decisão evidenciam que o processo decisório empresarial é complexo e que mesmo gestores experientes tomam decisões equivocadas, mas que as variáveis elencadas se relacionaram de maneira que é possível estabelecer um padrão entre o nível de determinada variável e a resposta correta.

As evidências de validade de conteúdo e construto apresentadas neste trabalho podem ser observadas pelas correlações entre as variáveis e a resposta, bem como a na diferença entre respostas racionais ou intuitivas (acertos e erros em cada uma das questões). Tais respostas convergem com as teorias que serviram de base para a construção dos cenários, de forma que os resultados também corroboram com os conceitos envolvendo decisões com ou sem viés.

Outra evidencia diz respeito ao índice de validade de conteúdo apurado em 97% como resultado da avaliação de juízes, o que aponta que o instrumento tem boa qualidade neste quesito, conforme sugerem Alexandre e Coluci (2011).

A manipulação dos cenários que resultou em diferenças quanto a quantidade de acertos, bem como a discriminação em classes são consideradas como evidências de validade de construto que se alinham as sugestões de Baguaei e Yazdi, (2016), uma vez

que de acordo com os pesquisadores, o nível de observação da variável latente pode se alterar à medida que se altera ou manipula o instrumento.

Os resultados observados na comparação das variáveis presentes no modelo auxiliam a interpretar os padrões de decisão. A hipótese de que traços de personalidade podem exercer alguma influência nas decisões foram observadas por meio da correlação com o cenário. Tal situação indica que o instrumento se relacionou de maneira complexa com fatores externos utilizados nas análises deste trabalho e referendados em outras publicações como estatisticamente significantes. Tal relacionamento condiz com a proposta da rede nomológica (Crombach & Meekl, 1955; Campbel & Fiske, 1959; Brown, 1970; Pasquali, 2007; Pais-Ribeiro, 2013; Viana, 2015) que sugere a existência de indícios de validade (como por exemplo validade convergente, critério e discriminante) resultante das relações apresentadas entre os construtos.

As hipóteses de que traços de personalidade possuem alguma relação com as decisões se faz presente neste estudo. Ao que indicam os resultados desta pesquisa, sustentam-se indícios de que a susceptibilidade heurística possui relação com traços de personalidade (Mcelroy & Dowd, 2007; Furnham & Boo, 2011; Caputo, 2014; Teovanovic, 2019).

As limitações deste estudo dizem respeito a amostra específica de gestores de uma única organização e área, o que pode sugerir que um viés quanto ao nível de conhecimento e aplicabilidade a realidade laboral do indivíduo tenha algum impacto nos resultados do teste. Também, a quantidade de indivíduos em cada grupo pode ser considerada relativamente baixa para análises mais complexas ou para que sejam feitas inferências mais complexas em relação a população. Uma quantidade maior de indivíduos seria desejável para análises ainda mais exploratórias. Também há clara lacuna na medição em relação ao tempo que cada indivíduo tenha dispensado para a solução dos cenários, visto que não houve medição controlada desta variável. O fator tempo pode ou não ser uma variável importante, mas ficou prejudicada neste contexto.

Estudos futuros podem buscar elucidar melhor o contexto destes cenários em demais níveis de gestão ou até mesmo a níveis operacionais de indivíduos que tenham atividades similares, como em recrutamento e seleção. A classificação de acordo com características e perfil resultante das respostas pode ampliar metodologias de seleção de pessoal com base nas classes, bem como identificar lacunas para desenvolvimento de habilidades, a exemplo do proposto por Kang e Park (2019) sobre investir em treinamentos que desenvolvam a mentalidade analítica dos indivíduos, como forma de incentivar que decisões sejam mais racionais e orientadas a estratégias. Também, como se observou a influência dos termos utilizados em cada cenário, há grandes possibilidades de embasar estudos na área de comunicação empresarial e disseminação de estratégia corporativa, considerando as classes oriundos de experimentos e a utilização de termos com maior probabilidade de ancoragem resultantes das observações de cenários como os propostos.

Referencias

- Alexandre, N. M. C., & Coluci, M., Z., O. (2011). Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Ciência & Saúde Coletiva*, 16(7), 3061-3068. <https://dx.doi.org/10.1590/S1413-81232011000800006>
- Antonczyk, R. C., & Salzmann, A. J. (2014). Overconfidence and optimism: The effect of national culture on capital structure. *Research in International Business and Finance*, 31, 132–151. <https://doi.org/10.1016/j.ribaf.2013.06.005>
- Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent Arbitrariness": Stable demand curves without stable preferences. *Quarterly Journal of Economics*, 118, 73-105.
- Baghaei, P., & Yazdi, M.T. (2016). The Logic of Latent Variable Analysis as Validity Evidence in Psychological Measurement. *The Open Psychology Journal*, 9, 168-175.
- Bergman, M. E., Drasgow, F., Donovan, M. A., Henning, J. B., & Juraska, S. E. (2006). Scoring Situational Judgment Tests: Once You Get the Data, Your Troubles Begin. *International Journal of Selection and Assessment*, 14(3), 223–235. <https://doi.org/10.1111/j.1468-2389.2006.00345.x>

- Bernardin, H. J., and Beatty, R. W., (1984), *Performance Appraisal: Assessing Human Behavior at Work*, Boston, Kent.
- Bol, J. C., & Smith, S. D., (2011), 'Spillover Effect in Subjective Performance Evaluation: Bias and the Asymmetric Influence of Controllability', *The Accounting Review*, 86(4), 1213-1230.
- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Adaptação e Validação de Instrumentos Psicológicos entre Culturas: Algumas Considerações. *Paidéia*, 22(53), 423–432. <https://doi.org/10.1590/S0103-863X2012000300014>
- Bowditch, J. L., & Buono, A. F., (2001) *A Primer on Organizational Behavior*. Fifth Edition, John Wiley & Sons, New York.
- Brown, F. G. (1970) *Principles of Educational and Psychological testing*. The Dryden Press Inc. Hinsdale. Illinois
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105
- Caputo, A. (2013). Relevant information, personality traits and anchoring effect. *International Journal of Management and Decision Making*, 13(1), 62. <https://doi.org/10.1504/ijmdm.2014.058470>
- Celeux, G., & Soromenho, G. (1996). An Entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*. 13. 195-212. 10.1007/BF01246098.
- Chantarangsi, W., Liu, W., Bretz, F. & Kiatsupaibul, S., & Hayter, A.. (2016). Normal Probability Plots with Confidence for the Residuals in Linear Regression. *Communications in Statistics - Simulation and Computation*. 47. 10.1080/03610918.2016.1165840.
- Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology*, 10, 519-540.

- Cheek, N. N., & Norem, J. K. (2017). Holistic thinkers anchor less: Exploring the roles of self construal and thinking styles in anchoring susceptibility. *Personality and Individual Differences, 115*, 174–176. <https://doi.org/10.1016/j.paid.2016.01.034>
- Choi, I., Koo, M., & An, J. (2007). Individual Differences in Analytic Versus Holistic Thinking. *Personality and Social Psychology Bulletin, 33*(5), 691–705. <https://doi.org/10.1177/0146167206298568>
- Cohen, H. B. (2019). An inconvenient truth about leadership development. *Organizational Dynamics, 48*(1), 8–15. <https://doi.org/10.1016/j.orgdyn.2017.10.002>
- Collins, L. M., & Lanza, S., T. (2010). *Latent class and latent transition analysis: with applications in the social behavioral, and health sciences. Wiley series in probability and statistics*. Hoboken, N.J: Wiley.
- Colombo, M., Bucher, L., & Inbar, Y. (2016). Explanatory Judgment, Moral Offense and Value Free Science. *Review of Philosophy and Psychology, 7*(4), 743–763. <https://doi.org/10.1007/s13164-015-0282-z>
- Cooper, W. H., (1981), 'Ubiquitous Halo', *Psychological Bulletin, 90*, 218-244.
- Costa, D., & Carvalho, F., & Moreira, B., C., & Silva, W. (2020). Viés de confirmação na tomada de decisão gerencial: um estudo experimental com gestores e contadores. *Revista de Contabilidade e Organizações. 14*. 164200. [10.11606/issn.1982-6486.rco.2020.164200](https://doi.org/10.11606/issn.1982-6486.rco.2020.164200).
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302
- Cruz, E., Fabio, Siqueira, Nascimento, & Silva, Da. (2008). Ciência e Pseudociência na Administração. *Revista Pensamento Contemporâneo em Administração. 2*. [10.12712/rpca.v2i2.131](https://doi.org/10.12712/rpca.v2i2.131).
- Czerwonka, M. (2017). Anchoring and Overconfidence: The Influence of Culture and Cognitive Abilities. *International Journal of Management and Economics, 53*(3), 48–66. <https://doi.org/10.1515/ijme-2017-0018>

- Ernst, A. F., Hoekstra, R., Wagenmakers, E. J., Gelman, A., & Van Ravenzwaaij, D. (2018). Do Researchers Anchor Their Beliefs on the Outcome of an Initial Study?: Testing the Time Reversal Heuristic. *Experimental Psychology*, 65(3), 158–169.
<https://doi.org/10.1027/1618-3169/a000402>
- Evans, J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual review of psychology*, 59, 255–278.
<https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223–241.
<https://doi.org/10.1177/1745691612460685>
- Farias, M., Newheiser, A. K., Kahane, G., & de Toledo, Z. (2013). Scientific faith: Belief in science increases in the face of stress and existential anxiety. *Journal of Experimental Social Psychology*, 49(6), 1210–1213.
<https://doi.org/10.1016/j.jesp.2013.05.008>
- Federação Brasileira de Bancos-Febraban (2018) *Pesquisa FEBRABAN de Tecnologia Bancária2018* <https://cmsportal.febraban.org.br/Arquivos/documentos/PDF/Pesquisa-FEBRABAN-Tecnologia-Bancaria-2019.pdf>
- Fernandes, D., Lynch, J. G., Jr., & Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8), 1861–1883. <https://doi.org/10.1287/mnsc.2013.1849>
- Fernandes, D., Lynch, J., G & Netemeyer, R. G. (2014) Financial literacy, financial education, and downstream financial behaviors. *Journal of Management*, Vol21.6, pp.868-97
- Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophesy fails*. Minneapolis: University of Minnesota Press
- Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *Journal of Socio Economics*, 40(1), 35–42. <https://doi.org/10.1016/j.socec.2010.10.008>

- Gava., A. M., & Vieira, K. M., (2013). Tomada de decisão em ambiente de risco: Uma avaliação sob a ótica comportamental. *Revista Eletrônica de Administração*, 12(1). Recuperado de <https://seer.ufrgs.br/read/article/view/40374/25652>
- Gervais, W. M., & Norenzayan, A. (2012). Analytic thinking promotes religious disbelief. *Science*, 336,493–496.
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Gonçalves, M., C., M., Schelini, P., W., & Deffendi, L., T. (2016). A relação entre extroversão e criatividade: um estudo com universitários brasileiros. *Boletim de Psicologia*, 66(145), 171-186.
- Grejo, L., M., Faia, V.S., Abbas, K. (2015) Avaliação do efeito Sunk Cost na tomada de decisão de acadêmicos de Ciências Contábeis e Administração. *Revista de Administração, Contabilidade e Economia da Fundace*. v. 6, n. 2, p. 64-79
- Hair, Jr., J.F., Willian, B., Babin, B.& Anderson, R.E. (2009). *Análise multivariada de dados*. 6.ed. Porto Alegre: Bookman.
- Hauck, F. N. (2015). Análise de classes latentes: uma breve introdução e ilustração metodológica. In: *Aplicações de métodos estatísticos avançados à avaliação psicológica e educacional*. 1ed.São Paulo: Vetor, v. , p. 193-210.
- Hitt, M. and Tyler, B. (1991) Strategic Decision Models: Integrating Different Perspectives. *Strategic Management Journal*, 12, 327-351. <https://doi.org/10.1002/smj.4250120502>
- Hutz, C. S. & Nunes, C. H. S. S. (2001). *Escala Fatorial de Neuroticismo*. São Paulo: Casa do Psicólogo.
- Jasper, F., & Ortner, T. M. (2014). The tendency to fall for distracting information while making judgments: development and validation of the Objective Heuristic Thinking Test. *European Journal of Psychological Assessment*, 30, 193–207.

- Joel, S., Spielmann, S. S., & MacDonald, G. (2017). Motivated Use of Numerical Anchors for Judgments Relevant to the Self. *Personality and Social Psychology Bulletin*, 43(7), 972-985. <https://doi.org/10.1177/0146167217702613>
- Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47:2, 263-91
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kahneman, D. (2012). *Rápido e devagar, duas formas de pensar*. Rio de Janeiro: Editora Objetiva.
- Kang, M., & Park, M. J. (2019). Employees' judgment and decision making in the banking industry. *International Journal of Bank Marketing*, 37(1), 382-400. <https://doi.org/10.1108/ijbm-04-2018-0111>
- Katopol, P. F. (2018). The Halo Effect and Bounded Rationality – Limits on Decision-Making. *Library Leadership & Management*, Volume 32(Number 3), 1-5.
- Kaustia, M., Alho, E., & Puttonen, V. (2008). How much does expertise reduce behavioral biases? The case of anchoring effects in stock return estimates. *Financial Management*. 37, 391-411.
- Kay, A. C., Gaucher, D., McGregor, I., & Nash, K. (2010). Religious belief as compensatory control. *Personality and Social Psychology Review*, 14, 37-48.
- Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J., & Laurin, K. (2008). God and the government: testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology*, 95, 18-35.
- Kimura, H., Basso, L. F. C., & Krauter, E. (2006). Paradoxos em finanças: teoria moderna versus finanças comportamentais. *Revista de Administração de Empresas*, 46(1), 41-58. doi: 10.1590/S0034-75902006000100005
- Knudsen, M. (2017). From Evidence-based Management to Management of Non-knowledge. *Emergence: Complexity and Organization*, 19 (2) . <https://doi.org/10.emerg/10.17357.f54a7c6d328d513f9c8d6be522b101bd>

- Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*, 37. 10.1108/00483480810877598.
- Marks, J., Copland, E., Loh, E., Sunstein, C. R., Sharot, T. (2018). Epistemic spillovers: Learning others' political views reduces the ability to assess and use their expertise in nonpolitical domains. *Harvard Public Law working paper* no. 18-22, Harvard University, Cambridge, MA. Retrieved from <https://ssrn.com/abstract=3162009>.
- Masyn, K. E., Henderson, C. E. & Greenbaum, P.E. (2010) Exploring the Latent Structure of Psychological Constructs in Social Development Using the Dimensional-Categorical Spectrum. *Social Development*, 19(3), 470-493. doi:10.1111/j.1467-9507.2009.00573.x
- Mcafee, R. P., Mialon, H., M. & Mialon, S., H. (2010) Do sunk costs matter? *Economic Inquiry*, vol. 48, n. 2, p. 323–336
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. III. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, 60(1), 63–91. <https://doi.org/10.1111/j.1744-6570.2007.00065.x>
- Mcelroy, T., & Dowd, K. (2007). Susceptibility to anchoring effects: How openness-to experience influences responses to anchoring cues. *Judgment and Decision Making*, 2(1), 48–53.
- Miot, H. A. (2017). Avaliação da normalidade dos dados em estudos clínicos e experimentais. *Jornal Vascular Brasileiro*, 16(2), 88-91
. <https://dx.doi.org/10.1590/1677-5449.041117>
- Motowildo, S. J., & Beimer, M. E. (2010) Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgement teste. *Journal of Applied Psychology*, 95(2):321-333
- Motowildoo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*, 75, 640–647

- Mowbray, F. I., Fox-Wasylyshyn, S. M., & El-Masri, M. M. (2019). Univariate Outliers: A Conceptual Overview for the Nurse Researcher. *Canadian Journal of Nursing Research*, 51(1), 31–37. <https://doi.org/10.1177/0844562118786647>
- Murphy, K. R., & Jako, R. A., (1989), 'Under What Condition are Observed Intercorrelations Greater than or Smaller than True Intercorrelations?', *Journal of Applied Psychology*, 74, 827-830.
- Mussweiler, T., & Schneller, K. (2003). "What goes up must come down" – How charts influence decisions to buy and sell stocks. *Journal of Behavioral Finance*, 4, 121-130.
- Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, 35, 136-164
- Muthén, L. K., & Muthén, B. O. (2014). *Mplus user's guide*. (Seventh Ed.). Los Angeles: Muthén & Muthén.
- Nechita, F., Alexandru, & Turcu-Stiolica, R. & Nechita, D. (2015). Original Paper The Influence of Personality Factors and Stress on Academic Performance. *Current Health Sciences Journal*. 41. 47-61. 10.12865/CHSJ.41.01.07.
- Niza, J. B., Marques, L., G. & Ferreira, M., B. (2015). Cognição social fora do laboratório não é peixe fora de água: O caso do efeito de ancoragem. *PSICOLOGIA*, 26(2), 141. <https://doi.org/10.17575/rpsicol.v26i2.274> No. 1, pp. 63-91.
- Nunes, C. H., & Hutz, C. S. (2007a). *Escala Fatorial de Extroversão - EFEx: Manual de aplicação*. São Paulo: Casapsi Livraria, Editora e Gráfica Ltda
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535–569. <https://doi.org/10.1080/10705510701575396>
- Olson-Buchanan, J. B., Drasgow, F., Moberg, P. J., Mead, A. D., Keenan, P. A., & Donovan, M. A. (1998). Interactive video assessment of conflict resolution skills. *Personnel Psychology*, 51(1), 1–24. <https://doi.org/10.1111/j.1744-6570.1998.tb00714.x>

- Pais-Ribeiro, J., L. (2013). Medida na avaliação psicológica. *Psicologia, Saúde & Doenças*, 14(1), 245-263.
- Paiva, R. T., Silva, H. A., Souza, J. C. M. de, Novôa, N. F., & Pereira, C. M. M. de A. (2020). O perfil do investidor individual no mercado financeiro. *Revista Vianna Sapiens*, 11(2), 30. <https://doi.org/10.31994/rvs.v11i2.694>
- Pasquali, L. (2007). Validade dos testes psicológicos: será possível reencontrar o caminho?. *Psicologia: Teoria e Pesquisa*, 23(spe), 99-107. <https://doi.org/10.1590/S0102-37722007000500019>
- Patterson, F., Galbraith, K., Flaxman, C., & Kirkpatrick, C. (2019). Evaluation of a Situational Judgement Test to Develop Non-Academic Skills in Pharmacy Students. *American journal of pharmaceutical education*, 83(10), 7074. <https://doi.org/10.5688/ajpe7074>
- Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2016). Atheists and agnostics are more reflective than religious believers: Four empirical studies and a meta analysis. *PLoS ONE*, 11(4). <https://doi.org/10.1371/journal.pone.0153039>
- Pervin, L. A. & John, O. P. (2004). *Personalidade: Teoria e pesquisa*. Trad. Ronaldo Cataldo Costa, 8ª ed. Porto Alegre: Artmed
- Piazza, J., & Sousa, P. (2014). Religiosity, political orientation, and consequentialist moral thinking. *Social Psychological and Personality Science*, 5, 334-342
- Pino, F. A. (2014) A questão da não normalidade: uma revisão. *Revista de Economia Agrícola*, São Paulo, v. 61, n. 2. p. 17-33.
- Ployhart, R. E. (2006). Staffing in the 21st Century: New Challenges and Strategic Opportunities. *Journal of Management*, 32(6), 868–897. <https://doi.org/10.1177/0149206306293625>
- Preston, J., & Epley, N. (2009). Science and God: an automatic opposition between ultimate explanations. *Journal of Experimental Social Psychology*, 45, 238–241.
- Proctor, R., N., & Schiebinger, L. (2008). *Agnotology: The making and unmaking of ignorance*. Stanford, CA: Stanford University Press

- Ramaswamy, V., Desarbo, W., Reibstein, D. & Robinson, W. (1993). An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with PIMS Data. *Marketing Science*, 12. 103-124. 10.1287/mksc.12.1.103.
- Rodrigues, M., R., M., C., Freire, T., M., Silva, J., D., G.(2016). Uma análise exploratória acerca da influência dos vieses dos custos afundados, do efeito house-money e da aversão ao arrependimento no comportamento dos empreendedores. *Revista de Gestão, Finanças e Contabilidade*, Salvador, v. 6, n. 3, p. 76-97
- Rutjens, B. T., Sutton, R. M., & van der Lee, R. (2018). Not All Skepticism Is Equal: Exploring the Ideological Antecedents of Science Acceptance and Rejection. *Personality and Social Psychology Bulletin*, 44(3), 384–405.
<https://doi.org/10.1177/0146167217741314>
- Rutjens, B. T., van der Pligt, J., & van Harreveld, F. (2010). Deus or Darwin: Randomness and belief in theories about the origin of life. *Journal of Experimental Social Psychology*, 46(6), 1078–1080. <https://doi.org/10.1016/j.jesp.2010.07.009>
- Rutjens, B. T., van Harreveld, F. & van der Pligt, J. (2010). Yes we can: Belief in progress as compensatory control. *Social Psychological and Personality Science*, 1, 246–252.
- Rynes, S. L., Colbert, A. E., & O’Boyle, E. H. (2018, November 1). When the “Best Available Evidence” Doesn’t Win: How Doubts About Science and Scientists Threaten the Future of Evidence-Based Management. *Journal of Management*, 44(8), 2995–3010.
<https://doi.org/10.1177/0149206318796934>
- Schmitt, N., & Chan, D. (2006). *Situational Judgment Tests: Method or Construct?* In J. A. Weekley & R. E. Ployhart (Eds.), *SIOP organizational series. Situational judgment tests: Theory, measurement, and application* (p. 135–155). Lawrence Erlbaum Associates Publishers.
- Schreiber, D., & Bohnenberger, M. C. (2017). A Racionalidade Limitada e a Percepção Seletiva no Processo Decisório na Área de P&D . *Revista de Gestão e Projetos*, 8(2), 58-70.

- Schreurs, S., Cleutjens, K., Collares, C. F., Cleland, J., & Oude Egbrink, M. (2020). Opening the black box of selection. *Advances in health sciences education : theory and practice*, 25(2), 363–382. <https://doi.org/10.1007/s10459-019-09925-1>
- Sedikides, C. (2010). Why does religiosity persist? *Personality and Social Psychology Review*, 14(1), 3–6. <https://doi.org/10.1177/1088868309352323>
- Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, 141(3), 423-428. <http://dx.doi-org.ez54.periodicos.capes.gov.br/10.1037/a0025391>
- Silva, R. F. M., Lagioia, U. C. T., Maciel, C. V., & Rodrigues, R. N. (2009). Finanças comportamentais: um estudo comparativo utilizando a teoria dos prospectos com os alunos de graduação do curso de ciências contábeis. *Revista Brasileira de Gestão de Negócios*, 11(33),383-403.
- Simon, H., A. (1955) "A behavioral model of rational choice", *The Quarterly Journal of Economics*, vol. 69, n. 1, February: 99-118, compiled in, and quoted from, Simon (1957: 241-260).
- Simon, H. (1991). Bounded Rationality and Organizational Learning. *Organization Science*, 2(1), 125-134. Retrieved March 22, 2021, from <http://www.jstor.org/stable/2634943>
- Slooman, S. A. (1996), "The empirical case for two systems of reasoning", *Psychological Bulletin*, v. 119, pp. 3-22.
- Soares, Henrique Fonseca Genelhu, & Barbedo, Claudio Henrique da Silveira. (2013). Desempenho acadêmico e a Teoria do Prospecto: estudo empírico sobre o comportamento decisório. *Revista de Administração Contemporânea*, 17(1), 64-82. <https://doi.org/10.1590/S1415-65552013000100005>
- Stanovich, K. E. (1999), *Who Is Rational?: Studies of individual Differences in Reasoning*. New Jersey, Lawrence Erlbaum.
- Stanovich, K. E., West, R. F. (2000), "Advancing the rationality debate", *Behavioral and Brain Sciences*, v. 23, n. 5, pp. 701-726.

- Starmer, C. (2000). Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, vol. XXXVIII, jun.2000, 332-382.
- Storrs, K. R. (2015). Are high-level aftereffects perceptual? *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00157>
- Sweis, B. M., Abram, S. V., Schmidt, B. J., Seeland, K. D., MacDonald, A. W., Thomas, M. J., & Redish, A. D. (2018). Sensitivity to “sunk costs” in mice, rats, and humans. *Science*, 361(6398), 178-181.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Talamas, S. N., Mavor, K. I., & Perrett, D. I. (2016). The influence of intelligence on the endorsement of the intelligence-attractiveness halo. *Personality and Individual Differences*, 95, 162–167. <https://doi.org/10.1016/j.paid.2016.02.053>
- Talhelm, T., Haidt, J., Oishi, S., Zhang, X., Miao, F. F., & Chen, S. (2015). Liberals Think More Analytically (More “WEIRD”) Than Conservatives. *Personality and Social Psychology Bulletin*, 41(2), 250–267. <https://doi.org/10.1177/0146167214563672>
- Teovanović, P. (2019). Individual differences in anchoring effect: Evidence for the role of insufficient adjustment. *Europe’s Journal of Psychology*, 15(1), 8–24. <https://doi.org/10.5964/ejop.v15i1.1691>
- Teovanović, P., Knežević, G., & Stankov, L. (2015). Individual differences in cognitive biases: Evidence against one-factor theory of rationality. *Intelligence*, 50, 75–86. <https://doi.org/10.1016/j.intell.2015.02.008>
- Thaler, R. H. (1980) Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, v. 1, n. 1, p. 39-60.
- Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12, 183-206.
- Thaler, R., & Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy*, 89(2), 392-406

- Tonetto, L., M., Kalil, L., L., Melo, W., Vi., Schneider, D., G., & Stein, L., M. (2006). O papel das heurísticas no julgamento e na tomada de decisão sob incerteza. *Estudos de Psicologia (Campinas)*, 23(2), 181-189. <https://dx.doi.org/10.1590/S0103-166X2006000200008>
- Tort-Martorell, X., Grima, P., & Marco, L. (2011). Management by facts: The common ground between total quality management and evidence-based management. *Total Quality Management and Business Excellence*, 22(6), 599–618. <https://doi.org/10.1080/14783363.2011.568262>
- Turner, B. M., & Schley, D. R. (2016). The anchor integration model: A descriptive model of anchoring effects. *Cognitive Psychology*, 90, 1–47. <https://doi.org/10.1016/j.cogpsych.2016.07.003>
- Tversky, A., & Kahneman, D (1981). The Framing Of Decisions And The Psychology Of Choice. *Science, New Series*, Vol. 211, No. 4481, Jan, 30, Pp. 453-458.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207-232. [http://dx.doi.org/10.1016/0010-0285\(73\)90033-9](http://dx.doi.org/10.1016/0010-0285(73)90033-9)
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131. <http://dx.doi.org/10.1126/science.185.4157.1124>
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293-315. <http://dx.doi.org/10.1037/0033-295X.90.4.293>
- Vianna, H. (2015). Validade de construto em testes educacionais. *Estudos em Avaliação Educacional*. 24. 136. [10.18222/eae246020143309](https://doi.org/10.18222/eae246020143309).
- Vlassova, A., & Pearson, J. (2013). Look Before You Leap: Sensory Memory Improves Decision Making. *Psychological Science*, 24(9), 1635–1643. <https://doi.org/10.1177/0956797612474321>
- Vol. 51 No. 1, pp. 1-24.

- Von Neumann, J., & Morgenstern, Oskar. (1944) *Theory of games and economic behavior*.
New Jersey: Princeton University Press
- Webster, E. S., Paton, L. W., Crampton, P. E. S., & Tiffin, P. A. (2020) Situational judgement test validity for selection: A systematic review and meta-analysis. *Med Educ.* 2020; 54: 888- 902. <https://doi.org/10.1111/medu.14201>
- Weekley, J. A., & Ployhart, R. E. (2006). *SIOPI organizational series. Situational judgment tests: Theory, measurement, and application*. Lawrence Erlbaum Associates Publishers.
- Welsh, M. B., Delfabbro, P. H., Burns, N. R., & Begg, S. H. (2014). Individual differences in anchoring: Traits and experience. *Learning and Individual Differences*, 29, 131–140. <https://doi.org/10.1016/j.lindif.2013.01.002>
- Whetzel, D., & Mcdaniel, M. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*. 19. 188-202. [10.1016/j.hrmr.2009.03.007](https://doi.org/10.1016/j.hrmr.2009.03.007).
World Bank Econ Rev 31(3):611–630. <https://doi.org/10.1093/wber/lhx018>
- Wright, A. L., Zammuto, R. F., Liesch, P. W., Middleton, S., Hibbert, P., Burke, J., & Brazil, V. (2016). Evidence-based Management in Practice: Opening up the Decision Process, Decision-maker and Context. *British Journal of Management*, 27(1), 161–178. <https://doi.org/10.1111/1467-8551.12123>
- Wuttke, E., Siegfried, C., & Aprea, C. (2020). Measuring financial literacy with a Situational Judgement Test: do some groups really perform worse or is it the measuring instrument. *Empirical Research in Vocational Education and Training* 12, 18 (2020). <https://doi.org/10.1186/s40461-020-00103-x>.
- Yang, C., Sun, B., & Shanks, D. R. (2018). The anchoring effect in metamemory monitoring. *Memory and Cognition*, 46(3), 384–397. <https://doi.org/10.3758/s13421017-0772-6>

- Yoshinaga, C., E, & Ramalho, T. B. (2014). Finanças Comportamentais no Brasil: uma aplicação da teoria da perspectiva em potenciais investidores. *Revista Brasileira de Gestão de Negócios*, 16(53), 594-615. <https://dx.doi.org/10.7819/rbgn.v16i52.1865>
- Yustina, A. I., & Gudono, G. (2017). Halo Effect in Subjective Performance Evaluation Bias. *Journal of Economics, Business & Accountancy Ventura*, 19(3).
<https://doi.org/10.14414/jebav.v19i3.621>

Apêndice C – Tesstes Situacionais de Julgamento e Decisão

FINANÇAS

FINAN1A

A empresa ABC é especializada em serviços de tecnologia e os sócios têm feito escolhas relativamente **boas** nos últimos 2 anos. Somando estes resultados que já chegam a quase R\$ 1 milhão, observa-se que um dos motivos deste cenário é a diversificação dos ativos. Os sócios se reúnem para que você avalie um novo investimento que foi prospectado no mercado e apresenta duas opções:

- a) chance de 90% de obter um ganho de 100.000
- b) chance de 92% de obter um ganho de 90.000

A empresa ABC é especializada em serviços de tecnologia e os sócios têm feito escolhas relativamente **ruins** nos últimos 2 anos. Somando estes resultados que já chegam a quase R\$ 1 milhão, observa-se que um dos motivos deste cenário é a diversificação dos ativos. Os sócios se reúnem para que você avalie um novo investimento que foi prospectado no mercado e apresenta duas opções:

- a) chance de 90% de obter um ganho de 100.000
- b) chance de 92% de obter um ganho de 90.000

FINAN1B

A empresa ABC é especializada em serviços de tecnologia e os sócios têm feito escolhas relativamente **boas** nos últimos 2 anos. Somando estes resultados que já chegam a quase R\$ 1 milhão, observa-se que um dos motivos deste cenário é a diversificação dos ativos. Os sócios se reúnem para que você avalie um novo investimento que foi prospectado no mercado e apresenta duas opções:

- a) chance de 10% de perda de 10.000 e 90% de ganho de 100.000
- b) chance de 50% de perda de 100.000 ou 50% de ganho de 150.000

A empresa ABC é especializada em serviços de tecnologia e os sócios têm feito escolhas relativamente **ruins** nos últimos 2 anos. Somando estes resultados que já chegam a quase R\$ 1 milhão, observa-se que um dos motivos deste cenário é a diversificação dos ativos. Os sócios se reúnem para que você avalie um novo investimento que foi prospectado no mercado e apresenta duas opções:

- a) chance de 10% de perda de 10.000 e 90% de ganho de 100.000
- b) chance de 50% de perda de 100.000 ou 50% de ganho de 150.000

FINAN2

A construtora AAA gostaria de diversificar seus investimentos e melhorar sua imagem do mercado. Dos 5 sócios, dois estão sendo investigados por envolvimento em práticas de lavagem de dinheiro e dos outros três restantes, 1 foi investigado recentemente, mas nada foi provado e para os demais não há nenhum processo ou citação.

Para este ano, foram elencadas 3 opções projetos. Qual você escolheria?

- a) Uma igreja, cuja projeção de lucros é de 300.000
- b) Uma agência de carros cuja projeção de lucros é de 400.000
- c) Uma vila com 4 casas cuja projeção de lucro é de 250.000

FINAN3

No ano passado, dois projetos foram escolhidos para o planejamento estratégico da unidade. O projeto escolhido é de sua autoria. Trata-se de um grande sonho pessoal e é compartilhado com a alta administração, que patrocina a ideia. Durante o este ano, o projeto já consumiu aproximadamente **90%** dos recursos alocados para sua execução. Até o momento, os resultados estão muito aquém do que foi planejado (**10%** de retorno, enquanto o esperado era 60%). Na orçamentação deste ano, será proposto um novo cenário de aporte de recursos na ordem de mais 50% do valor inicialmente previsto para o projeto, e que pode dar uma sobrevida a operação. Desta forma, você deve decidir:

- a) Alocar o montante no projeto atual

b) Abandonar o projeto atual e iniciar o segundo projeto

No ano passado, dois projetos foram escolhidos para o planejamento estratégico da unidade. O projeto escolhido inclusive, é de sua autoria. Trata-se de um grande sonho pessoal e é compartilhado com a alta administração, que patrocina a ideia. Durante o este ano, o projeto já consumiu aproximadamente **70%** dos recursos alocados para sua execução. Até o momento, os resultados estão muito aquém do que foi planejado (**30%** de retorno, enquanto o esperado era 60%). Na orçamentação deste ano, será proposto um novo cenário de aporte de recursos na ordem de mais 50% do valor inicialmente previsto para o projeto, e que pode dar uma sobrevida a operação. Desta forma, você deve decidir:

a) Alocar o montante no projeto atual

b) Abandonar o projeto atual e iniciar o segundo projeto

GESTÃO DE PESSOAS

PESS1A

Ariclenes sempre foi um funcionário que podemos definir como "desleixado". Apesar de trabalhar em um setor onde era exigida discrição e aderência a procedimentos de segurança, não era incomum que seu superior encontrasse algum erro eventual. Esquecia-se de realizar conferências, mandar e-mails e controlar alguns processos sob sua responsabilidade, mas quando cobrado, se desdobrava para realizar. Por sorte, nunca ocorreram prejuízos financeiros de grande vulto, mas em certas ocasiões, ele foi responsabilizado e teve que pagar **R\$ 10.000,00** decorrentes da sua distração. Em determinado momento, faltou ao serviço sem justificativa durante uma semana e foi demitido. Na sua opinião, a demissão foi correta?

a) Sim

b) Não

Ariclones sempre foi um funcionário que podemos definir como 'desleixado". Apesar de trabalhar em um setor onde era exigida discrição e aderência a procedimentos de segurança, não era incomum que seu superior encontrasse algum erro eventual. Esquecia-se de realizar conferências, mandar e-mails e controlar alguns processos sob sua responsabilidade, mas quando cobrado, se desdobrava para realizar. Por sorte, nunca ocorreram prejuízos financeiros de grande vulto, mas em certas ocasiões, ele foi responsabilizado e teve que pagar **R\$ 100.000,00** decorrentes da sua distração. Em determinado momento, faltou ao serviço sem justificativa durante uma semana e foi demitido. Na sua opinião, a demissão foi correta?

- a) Sim
- b) Não

PESS1B

Na mesma época, Atadolfo, funcionário de performance exemplar que inclusive recebeu bons elogios sobre a impecável condução dos processos sob sua responsabilidade, cometeu um erro terrível, deu um prejuízo de R\$ 100.000,00 e faltou por uma semana, sendo demitido em seu retorno. Na sua opinião, a demissão de Atadolfo foi correta?

- a) Sim
- b) Não

PESS1C

Se você pudesse decidir pela demissão de somente um dos dois, qual seria?

- a) Ariclones
- b) Atadolfo

PESS2

Você está encarregado de escolher uma pessoa com habilidades de liderança para preencher uma vaga de gerente em uma grande equipe. São necessárias habilidades

consideráveis neste tema, visto que o cenário é desafiador, a equipe já possui alguns problemas no passado, está em uma fase de baixa auto estima e pouco resultado.

a) Nero: Filho adotivo de uma família tradicional, grande administrador, adora música e teatro, possui grandes realizações na carreira, mas tem um gênio difícil.

b) Adolf: Estudioso, possui habilidade de liderança, vegetariano, não bebe, não fuma, escreveu um livro recentemente, mas é megalomaniaco.

c) Luther King: Boas habilidades de trato social, comprometido, adepto da comunicação não violenta, mediação, mas já perdeu a condição de réu primário.

PESS3

Surgiu uma vaga na equipe na qual você é gestor e é necessário preenchê-la. Caso não seja preenchida até as 12 horas de hoje, a vaga será cortada:

a) Adesbaldino - Já foi gestor, mas seu cargo foi cortado em um momento de contingenciamento orçamentário. No passado, um colega de trabalho falou bem dele, ressaltando as consideráveis qualidades de relacionamento interpessoal.

b) Esperidião - Chegou recentemente na equipe e promovido de um outro setor. Possui um ótimo currículo e gosta de aprender, mas é pouco conhecido em sua área. Também, não possui muita experiência no trabalho.

Considerações Finais

As pesquisas realizadas nesta dissertação auxiliam na melhoria das definições operacionais acerca de temas de grande relevância na atualidade: racionalidade, crença em ciência e julgamento e decisão. A abordagem do trabalho pode auxiliar novas pesquisas derivadas dos conceitos aqui presentes. Há que se pensar que os conceitos derivados de razão e intuição em ambiente organizacional ainda são constantemente cercados de conceitos populares, mas que nem sempre possuem evidências científicas. Por meio dos inúmeros indícios de validade observados nos manuscritos, aos poucos o obscurantismo acerca de certos temas vai decrescendo e subsidiará exploração de novas pesquisas no campo das decisões organizacionais.

Dentre os indícios de validade observados no que tange a medir a racionalidade, mais precisamente utilizando a medida decorrente do instrumento CRT e produto dos resultados observados na análise dos dados das amostras pesquisadas, observou-se que as qualidades psicométricas do instrumento não são tão boas como as observadas em pesquisas internacionais, como por exemplo as correlações encontradas com outros fatores, mas que o instrumento pode se tornar promissor se melhorado. Embora para o contexto brasileiro existam os indícios de que o teste mede o que se pressupõe a medir, há algumas ressalvas que podem derivar da quantidade diminuta de itens do instrumento. O que se observa é que em pesquisas internacionais existe certa frequência de utilização, mas ao mesmo tempo, há uma tendência em adicionar maior quantidade de itens na escala e que tal acréscimo têm elevado os indicadores de performance (confiabilidade, validade interna e externa) do instrumento.

A escala de crença na ciência mostrou-se um instrumento com boas qualidades psicométricas para o contexto brasileiro e o estudo apresenta diversos pontos de reflexão quanto aos consideráveis indícios de validade derivados das técnicas estatísticas aplicadas. Com este estudo, foi possível medir a crença na ciência dos brasileiros e em especial a relação entre a crença religiosa e a crença científica com elevado tamanho de efeito. Com os resultados expostos, abrem-se possibilidade de estudar a crença na ciência em múltiplas

áreas aplicadas e entender um pouco melhor como as pessoas lidam com a ciência na sociedade e nas organizações. Hipótese que ainda permanece é que a descrença científica esteja relacionada com o baixo letramento científico do brasileiro e com níveis de religiosidade mais elevados, relações estas que não foram possíveis de se observar.

Os testes de julgamento e decisão empresarial e suas consequentes evidências de validade sugerem que a classificação de gestores quanto aos seus padrões de resposta permitiria em ambiente organizacional melhorar a alocação de pessoas de acordo com o perfil, selecionar indivíduos com o traço latente requerido e fomentar o treinamento e desenvolvimento de formas de pensar mais analíticas. Outro ponto de grande importância é que as relações com variáveis externas e a influência da quantidade e qualidade das informações pode subsidiar estudos futuros sobre perfis psicológicos desejáveis para gestores em determinada área (por exemplo, necessidade de indivíduos com determinado traço de personalidade mais elevado ou mais baixo, por conta do tipo de informação tratada e que, de acordo com esse traço, o tornaria menos susceptível a efeitos heurísticos). Outro ponto de grande relevância para o estudo são os indícios de validade discutidos no manuscrito, de forma que os cenários apresentados buscam simular situações reais (diferente do padrão de outras pesquisas, que apresenta questões um pouco generalistas e com poucos indícios de validade) e corroboraram com conceitos observados na literatura multidisciplinar sobre julgamento e decisão.

Por fim, os conceitos envolvendo os três manuscritos acabam se consolidando no último manuscrito de maneira convergente e sugerem novas oportunidades de estudo aplicado as organizações.