

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
DEPARTAMENTO DE PSICOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOLOGIA COGNITIVA

MOZART AMORIM

**AVALIAÇÃO DO PENSAMENTO CRÍTICO: ANÁLISE DAS PROPRIEDADES
PSICOMÉTRICAS DE UM TESTE DE RACIOCÍNIO LÓGICO-DEDUTIVO**

**RECIFE
2015**

MOZART AMORIM

**AVALIAÇÃO DO PENSAMENTO CRÍTICO: ANÁLISE DAS PROPRIEDADES
PSICOMÉTRICAS DE UM TESTE DE RACIOCÍNIO LÓGICO-DEDUTIVO**

Dissertação desenvolvida pelo mestrando Mozart Amorim, do curso de pós-graduação em Psicologia Cognitiva da UFPE e a mesma cumpre as exigências do programa.

Área de concentração: Desenvolvimento cognitivo

Orientador: Dr. José Maurício Haas Bueno

**RECIFE
2015**

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na fonte
Bibliotecária, Divonete Tenório Ferraz Gominho. CRB4- 985

A524a Amorim, Mozart

Avaliação do pensamento crítico: análise das propriedades psicométricas de um teste de raciocínio lógico-dedutivo / Mozart Amorim. – Recife: O autor, 2015.

95 f. : il. ; 30 cm.

Orientador: Prof. Dr. José Maurício Haas Bueno.

Dissertação (mestrado) - Universidade Federal de Pernambuco. CFCH. Pós-Graduação em Psicologia Cognitiva, 2015.

Inclui referências e anexos.

1. Psicologia cognitiva. 2. Pensamento crítico. 3. (Temas em avaliação psicológica). 4. Lógica. I. Bueno, José Maurício Haas. (Orientador). II. Título.

153 CDD (22.ed.)

UFPE (BCFCH2015-28)



ATA DA 300ª DEFESA DE DISSERTAÇÃO DE MESTRADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOLOGIA COGNITIVA DO CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS DA UNIVERSIDADE FEDERAL DE PERNAMBUCO, NO DIA 23 DE FEVEREIRO DE 2015.

Aos **23** (vinte e três) dias do mês de **fevereiro de dois mil e quinze** (2015), às dezesseis horas e trinta minutos, na Sala de Aula 01 do 8º andar do Centro de Filosofia e Ciências Humanas-CFCH da Universidade Federal de Pernambuco, em sessão pública, teve início a defesa da Dissertação de Mestrado intitulada “Avaliação do Pensamento Crítico: análise das propriedades psicométricas de um teste de raciocínio lógico-dedutivo” do aluno **MOZART AMORIM SOUSA**, na área de concentração Psicologia Cognitiva, sob a orientação do Prof. Dr. José Maurício Haas Bueno. O mestrando cumpriu todos os demais requisitos regimentais para a obtenção do grau de MESTRE em Psicologia Cognitiva. A Banca Examinadora foi indicada pelo Colegiado do Programa de Pós-Graduação em Psicologia Cognitiva em *Ad Referendum* e homologada pela Diretoria de Pós-Graduação, através do Processo Nº 23076.005370/2015-40 em 05 (cinco) de fevereiro de dois mil e quinze (2015), composta pelos Professores Doutores JOSÉ MAURICIO HAAS BUENO (Presidente e 1º Orientador), FABIANO KOICH MIGUEL (Examinador Externo, Dep. de Fundamentos de Psicologia e Psicanálise/UEL - PR) e SINTRIA LABRES LAUTERT (Examinador Interno). Após cumpridas as formalidades, o candidato foi convidado a discorrer sobre o conteúdo da Dissertação. Concluída a explanação, o candidato foi arguido pela Banca Examinadora que, em seguida, reuniu-se para deliberar e conceder ao mesmo a menção **APROVADA** da referida Dissertação. E, para constar, lavrei a presente Ata que vai por mim assinada, Secretária de Pós-Graduação, e pelos membros da Banca Examinadora.

Recife, 23 de fevereiro de 2015

BANCA EXAMINADORA

PROF. JOSÉ MAURICIO HAAS BUENO

PROF. FABIANO KOICH MIGUEL

PROFA. SINTRIA LABRES LAUTERT

DEDICATÓRIA

Dedico este trabalho aos meus familiares, a Pandora e a Renata,
companheira fielmente constante nesta jornada.

AGRADECIMENTOS

Agradeço a Deus

por todas as oportunidades de estudo e desenvolvimento que esta vida me concedeu.

A todo o grupo de pesquisas do NEAP

Pelo auxílio constante e ininterrupto ao longo de todo o desenvolvimento da pesquisa.

Ao professor Maurício

Pela orientação e apoio durante a graduação e a pós-graduação.

Aos demais professores da Pós Graduação em Psicologia Cognitiva

Por todos os ensinamentos, aulas e debates no percurso deste mestrado.

Amorim, Mozart. (2015). Avaliação do pensamento crítico: análise das propriedades psicométricas de um teste de raciocínio lógico-dedutivo. Dissertação de mestrado, Pós-Graduação em Psicologia Cognitiva, Universidade Federal de Pernambuco, Recife.

RESUMO

A presente pesquisa teve como objetivo estudar as propriedades psicométricas de um Teste de Raciocínio Lógico-Dedutivo (TRLD). Para tanto, foram realizados dois estudos, sendo o primeiro pela Teoria Clássica dos Testes e o segundo pela Teoria de Resposta ao Item. Ambos contaram com uma amostra de 192 participantes, (65,6% pertencentes ao sexo feminino), todos com idade superior a 18 anos ($M=23,9$, $DP=9,33$). Foram aplicados um questionário sociodemográfico, o Teste de Raciocínio Lógico-Dedutivo, a Bateria Fatorial de Personalidade e a Prova de Raciocínio Abstrato da BPR-5. No primeiro estudo, foram investigadas a validade fatorial do TRLD, as precisões dos fatores encontrados e a validade convergente-discriminante, com inteligência fluída (raciocínio abstrato) e traços de personalidade, respectivamente. Uma análise fatorial exploratória revelou uma estrutura com cinco fatores primários e dois fatores de segunda ordem. Os índices de consistência interna foram de 0,844 e 0,708, para os fatores de segunda ordem, respectivamente. Além disso, observou-se correlação moderada e significativa desses fatores com uma medida de raciocínio abstrato (validade convergente) e correlações baixas ou não significativas com traços de personalidade (validade discriminante). O instrumento foi considerado válido e preciso, especialmente para avaliação dos dois fatores de segunda ordem. No segundo estudo os fatores de segunda ordem foram submetidos à Teoria de Resposta ao Item (TRI). Foram analisados o ajuste dos dados ao Modelo de Rasch da TRI (*infit* e *outfit*); a adequação do nível de dificuldade dos itens (parâmetro b) ao nível de habilidade dos sujeitos (Θ) e a adequação das respostas consideradas corretas. Em ambos os fatores de segunda ordem, os índices de *infit* e *outfit* indicaram um bom ajuste dos dados ao Modelo de Rasch e adequação das respostas consideradas corretas, que foram escolhidas pelos participantes com média de theta mais elevada. Além disso, os itens do primeiro fator de segunda ordem apresentaram nível de dificuldade compatível com a média de habilidade dos sujeitos, mas os itens do segundo fator de segunda ordem foram muito fáceis para o nível de habilidade dos sujeitos. Embora algumas melhorias possam ser realizadas em futuros estudos, no geral, pode-se dizer que o instrumento apresentou boas propriedades psicométricas, podendo ser recomendado para uso em pesquisas.

Palavras-chave: Pensamento crítico; avaliação psicológica; testes psicológicos; lógica informal.

Amorim, Mozart. (2015). Critical thinking evaluation: analysis of the psychometric properties based on the logical-deductive reasoning test. Master's degree thesis, Cognitive Psychology Graduate School, Universidade Federal de Pernambuco, Recife.

ABSTRACT

This research aimed to study the psychometric properties of the Logical-Deductive Reasoning Test (LDRT). Therefore, two studies were performed, the first one using the Classical Theory of Tests and the second one, the Item Response Theory. Both relied on a sample of 192 participants (65.6% females), all aged over 18 years old ($M = 23.9$, $SD = 9.33$). A sociodemographic questionnaire was applied, as also the Logical-Deductive Reasoning Test, the Personality Factor Battery and the BPR-5 Abstract Reasoning Test. In the first study, there were investigated the LDRT factorial validity, the identified factors' reliabilities and the convergent-discriminant validity, with fluid intelligence (abstract reasoning) and personality traits, respectively. An exploratory factor analysis revealed a structure containing five primary factors and two second-order factors. The results of internal consistency coefficients were 0.844 and 0.708 for the second-order factors, respectively. In addition, it were observed moderate and significant correlations in the measure of abstract reasoning (convergent validity) and low or not-significant correlations in personality traits' results (discriminant validity). The instrument was considered valid and accurate, especially for the two second-order factors evaluation. In the second study, the second-order factors were submitted to the Item Response Theory (IRT). There were analyzed the data adjustment to the IRT Rasch model (*infit* and *outfit*); the adequacy between the items' difficulties (parameter b) and the individuals' abilities (θ) and the adequacy of the correct answers. In both second-order factors, *infit* and *outfit* rates indicated a positive adjustment of the data to the Rasch Model, and the adequacy of the correct answers, chosen by the participants with higher level of θ . Furthermore, the difficulty of the items of the first second-order factor was compatible with the ability average of the participants, however the second factor items were lower than the ability average of the participants. Although some improvements can be implemented in future studies, in general, the instrument demonstrated good psychometric results and, consequently, can be recommended to be used in researches.

Keywords: critical thinking; psychological assessment; psychological tests; informal logic.

LISTA DE TABELAS

ARTIGO 1

Tabela 1 - Cargas fatoriais dos cinco principais fatores.....	54
Tabela 2 - Cargas fatoriais da Análise Fatorial de Segunda Ordem	57
Tabela 3 - Correlação de Pearson dos fatores de segunda ordem obtidos no teste de raciocínio lógico-dedutivo com a prova de Raciocínio Abstrato (RA) da BPR5 e com os traços de personalidade da bateria BFP (Neuroticismo, Extroversão, Socialização, Responsabilidade e Abertura à experiência).....	58
Tabela 4: Análise de Perfis de Medidas Repetidas	59

ARTIGO 2

Tabela 1 – Estatística descritiva dos itens do FSO1.....	78
Tabela 2 - Estatística descritiva dos itens do FSO2.....	81

LISTA DE SIGLAS E ABREVIATURAS

ASBRO	Associação Brasileira de Rorschach e Métodos Projetivos
BFP	Bateria Fatorial de Personalidade
BPR-5	Bateria de Provas de Raciocínio
Fator g	Fator geral
FSO1	Fator de Segunda Ordem 1
FSO2	Fator de Segunda Ordem 2
HCTAES	Halpern Critical Thinking Assessment Using Everyday Situations
IBAP	Instituto Brasileiro de Avaliação Psicológica
NuPArg	Núcleo de Pesquisas em Argumentação
QI	Quociente de Inteligência
RA	Raciocínio Abstrato
SATEPSI	Sistema de Avaliação dos Testes Psicológicos
TCLE	Termo de Consentimento Livre e Esclarecido
TCM	Teoria Clássica de Medidas
TCT	Teoria Clássica de Testes
TRI	Teoria de Resposta ao Item
TRLD	Teste de Raciocínio Lógico-Dedutivo
WGCTA	Watson-Glaser Critical Thinking Assessment

SUMÁRIO

APRESENTAÇÃO.....	11
INTRODUÇÃO.....	13
A psicometria e os modelos de análise estatística: TCM e TRI.....	15
O pensamento crítico e seus conceitos.....	24
O pensamento crítico e seus fatores constituintes.....	29
A avaliação do Pensamento Crítico.....	31
REFERÊNCIAS.....	38
ARTIGO 1: ESTUDOS DE VALIDADE DE UM TESTE DE RACIOCÍNIO LÓGICO-DEDUTIVO.....	44
RESUMO.....	44
ABSTRACT.....	45
INTRODUÇÃO.....	46
MÉTODO.....	50
Participantes.....	50
Instrumentos.....	51
Procedimentos.....	52
RESULTADOS.....	53
DISCUSSÃO DOS RESULTADOS.....	61
REFERÊNCIAS.....	63
ARTIGO 2: ANÁLISE DE UM TESTE DE RACIOCÍNIO LÓGICO-DEDUTIVO COM TRI.....	65
RESUMO.....	65
ABSTRACT.....	66
INTRODUÇÃO.....	67
MÉTODO.....	72
Participantes.....	72
Instrumentos.....	72
Procedimentos.....	72

RESULTADOS.....	75
DISCUSSÃO DOS RESULTADOS.....	82
REFERÊNCIAS.....	84
CONCLUSÃO GERAL.....	86
ANEXOS.....	88

APRESENTAÇÃO

Esta dissertação é composta por uma fundamentação teórica, onde são apresentadas a psicometria e os dois grandes modelos de análise que serão utilizados neste estudo. Após esta parte inicial foram abordados os aspectos cognitivos relacionados aos conceitos de inteligência, da lógica, da razão e por fim do pensamento crítico, abordando neste último os seus diversos conceitos, seus fatores constituintes bem como os principais instrumentos utilizados para sua avaliação.

Em relação ao instrumento utilizado nesta pesquisa, o teste de raciocínio lógico-dedutivo (Primi & Muller, não publicado) avalia a capacidade de análise de informações complexas existentes em uma situação-problema, buscando a realização de correlações para uma correta conclusão lógica. Este teste adota o modelo cognitivo da *Transitive Chain Theory*, de Guyote e Sternberg (1981), o qual propõe-se a entender os processos de raciocínio silogístico a partir de um percurso de pensamento e de linguagem, permitindo a dedução de algo. A teoria fala a respeito da combinação de regras no pensamento que formariam uma representação integrada de toda a informação existente no argumento.

Dessa maneira a pesquisa realizada pretendeu avaliar o pensamento crítico utilizando parâmetros psicométricos do teste de raciocínio lógico-dedutivo (Primi & Muller, não publicado). A importância do estudo do pensamento crítico se deve ao fato de que esta habilidade cognitiva, tomada aqui com o conceito de ação reflexiva e propositiva capaz de discriminar a informação relevante daquela que não é, no intuito de auxiliar a resolução de questões complexas, está presente no desenvolvimento escolar e é sobretudo exigida no âmbito da educação superior, permitindo maior autonomia do indivíduo diante das necessidades acadêmicas continuamente em desenvolvimento. Também é possível constatar que o grande excesso de informações produzidas e recebidas pelos indivíduos, principalmente através da internet, necessitam de uma seleção cada vez mais eficiente por parte do receptor, uma vez que não há possibilidade de assimilação pelo indivíduo de toda a informação produzida, desta maneira o pensamento crítico também atua auxiliando o indivíduo a filtrar as informações que serão úteis ao seu cotidiano.

Para análise dos dados foram utilizadas duas grandes teorias psicométricas: a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI). A dissertação foi então dividida

em duas partes, sendo a primeira focada nos objetivos utilizando a TCT enquanto a segunda parte se debruçou em análise com a TRI.

Utilizando a TCT o objetivo foi buscar evidências de validade baseada na estrutura interna e buscar evidências de validade com base nas relações com variáveis externas, onde foram utilizados instrumentos de coleta específica para o raciocínio abstrato e para os cinco grandes fatores de personalidade. Por fim a TCT foi também utilizada para investigar os índices de fidedignidade dos fatores encontrados na análise fatorial.

O estudo utilizando a Teoria de Resposta ao Item (TRI) pretendeu obter análise do ajuste dos dados ao Modelo de Rasch da TRI (*infit* e *outfit*), bem como analisar a adequação do nível de dificuldade dos itens (parâmetro b) ao nível de habilidade dos sujeitos θ (Θ). Finalizando, seria também objetivo com a TRI analisar o espectro de precisão do instrumento.

A pesquisa como um todo teve como amostra pessoas residentes na região metropolitana do Recife, nas quais foram aplicados os seguintes testes: teste de Raciocínio Lógico-Dedutivo (Primi & Muller, não publicado); aplicação da Prova de Raciocínio Abstrato (RA), Forma B, da Bateria de Provas de Raciocínio (BPR-5) para fins de coleta de dados acerca da inteligência fluida (capacidade de raciocínio diante de situações novas, originando novas definições e entendendo as consequências do raciocínio) e por fim a aplicação da Bateria Fatorial de Personalidade (BFP), a qual avalia a personalidade utilizando o modelo dos cinco grandes fatores (neuroticismo, extroversão, responsabilidade, socialização e abertura à experiência).

Os estudos e seus resultados serão apresentados a seguir no formato de dois artigos, sendo o primeiro intitulado *Estudos de validade de um teste de raciocínio lógico-dedutivo*, o qual apresenta as evidências de validade encontradas na análise estatística, baseadas na estrutura interna do teste de Raciocínio lógico dedutivo. Também apresenta a avaliação da precisão (fidedignidade) dos fatores que foram obtidos no estudo.

Já o segundo artigo, intitulado *Análise de um teste de raciocínio lógico-dedutivo com TRI*, versa sobre a investigação das propriedades psicométricas do teste de raciocínio lógico dedutivo por meio da Teoria de Resposta ao Item (TRI) com objetivo de analisar o ajuste dos dados ao Modelo de Rasch da TRI (*infit* e *outfit*), como também verificar a adequação do nível de dificuldade dos itens (parâmetro b) ao nível de habilidade dos sujeitos θ (Θ) e por fim discutir o espectro de precisão do instrumento.

INTRODUÇÃO

A área de testagem psicológica no Brasil passou por diversas etapas, superando dificuldades iniciais em relação à disponibilidade de instrumentos confiáveis bem como de publicação de pesquisas nacionais sobre o tema. Mas, é possível afirmar que atualmente essa área já apresenta grande desenvolvimento de grupos de pesquisa, metodologia bem estruturada tanto no campo quantitativo quanto no campo qualitativo e sempre posta à avaliação crítica e técnica em publicações de revistas, congressos e encontros nacionais, bem como intenção e preocupação crescente no que diz respeito à propagação científica da avaliação psicológica, porém observa-se que em termos de ensino as alterações curriculares atualmente focam numa diminuição da carga horária das disciplinas relacionadas ao conteúdo de técnicas e dos testes psicológicos em diversos cursos de graduação do Brasil.

Pesquisas realizadas por Noronha (2004, 2005) sobre a qualidade do ensino em avaliação psicológica no Brasil apontaram diversos obstáculos que ainda necessitam de superação para que o aluno de psicologia venha a dominar e conhecer melhor os atributos básicos deste campo. Em pesquisa realizada com 322 sujeitos alunos de dois cursos de psicologia (um no estado de São Paulo e outro do estado do Rio Grande do Sul), concluiu-se que os aspectos principais do ensino de avaliação psicológica por meio de testes necessitavam ser repensados e definidos com maior clareza nos currículos dos cursos de psicologia do país, uma vez que os testes psicológicos estão contemplados (conforme a lei 4119 da criação dos cursos de psicologia no Brasil) nas disciplinas de técnicas de exame psicológico, mas os conteúdos relativos aos testes psicológicos parecem não propiciar uma quantidade de informações básicas que se considere suficiente aos alunos de psicologia (Noronha & Alchieri, 2004).

Após realização de pesquisa sobre a produção científica brasileira relacionada aos testes psicológicos, Filho, Belo e Gouveia (2006) salientam que:

Neste contexto, observa-se que o século XX presenciou uma constante utilização dos testes, não apenas dos testes de aplicação individual, que eram mais demorados e demandavam maior trabalho, mas também os testes de aplicação coletiva. As duas grandes guerras mundiais colaboraram nesse sentido, uma vez que, nessas ocasiões, foram desenvolvidos testes direcionados para atender as necessidades da guerra, como escalas que pudessem rapidamente classificar milhares de recrutas em termos de suas aptidões e capacidades intelectuais. (p. 482)

Mesmo com vários obstáculos a serem superados, a produção científica nacional relacionada ao uso dos testes psicológicos foi crescente durante a maior parte do século XX, favorecendo assim o próprio desenvolvimento dos testes no Brasil, bem como a criação de novos tipos de avaliação psicológica. É verdade que os testes psicológicos fazem parte da própria história da psicologia, marcando inclusive a identidade profissional psicólogo em determinados países, mas assim como qualquer outro tipo de tecnologia, os testes psicológicos necessitam passar por aprimoramentos, buscando assim atender de maneira correta os critérios de qualidade exigidos pelos órgãos de classe e por instituições internacionais.

Alguns países, como os Estados Unidos, por exemplo, revelam ainda hoje extrema capacidade de organização do uso dos testes psicológicos, inclusive mantendo publicações de manuais de aplicação de testes desenvolvidos por diversas universidades americanas que trabalharam no intuito de padronizar ao máximo a aplicação destes testes. A Psychological Assessment Resources, criada há 25 anos, publicou um catálogo de testes que poderiam ser utilizados em diversas áreas, como por exemplo a escolar, judicial, organizacional, dentre outras.

No Brasil, apesar dos avanços ocorridos nos últimos vinte anos na área de avaliação psicológica com a criação de instituições, como por exemplo, o Sistema de Avaliação de Testes Psicológicos (SATEPSI) órgão inexistente nos Estados Unidos, a Associação Brasileira de Rorschach e Métodos Projetivos (ASBRO), o Instituto Brasileiro de Avaliação Psicológica (IBAP), dentre outros, ainda não há uma sistematização de publicações tão rigorosa como a americana, inclusive a própria formação em avaliação psicológica no Brasil ainda é superficial no nível de graduação, sendo necessários estudos de pós-graduação para um entendimento básico sobre os testes.

Noronha, Primi e Alchieri (2005) realizaram no Brasil uma pesquisa utilizando 304 sujeitos, sendo 223 estudantes de graduação de psicologia e 74 sujeitos já formados em psicologia e com diferentes graus de cursos de pós-graduação, concluindo que as técnicas de avaliação da personalidade constituem-se nos testes mais conhecidos, sendo inclusive usados em diferentes contextos da atuação do psicólogo. Contudo, como na relação da pesquisa havia 145 títulos de testes psicológicos, os pesquisadores concluíram que a maioria deles ainda era desconhecida por parte da classe de psicólogos, sendo menor ainda o número de testes utilizados por esse profissional. Assim os pesquisadores ressaltaram que a sistematização na área de avaliação deveria

ser uma prioridade para essa área científica, bem como se faz necessário que este profissional que manuseia os testes psicológicos mantenha-se sempre atualizado na aplicação instrumental e na interpretação de resultados (Noronha, Primi & Alchieri, 2005).

A psicometria e os modelos de análise estatística: TCM e TRI

O surgimento da psicometria e o crescente interesse pela busca do entendimento de determinadas faculdades humanas, nos finais do século XIX e início do XX, propiciou a construção e reformulação contínua de vários conceitos dentro da própria psicologia, sendo a inteligência um destes conceitos amplamente revisados e proposto a ser medido pelos mais distintos testes psicológicos já produzidos. Apesar das descobertas atuais sobre a existência de não apenas uma, mas de inteligências múltiplas existentes no ser humano, a sociedade ocidental historicamente supervalorizou o tipo de inteligência conhecida como linguística e a lógico-matemática, constatando-se isso através da formação curricular das escolas bem como na elaboração e desenvolvimento de testes psicológicos que só mediam tais tipos de inteligência, como por exemplo o teste de quociente de inteligência ou simplesmente QI. A grande aceitação do teste de QI aconteceu quando, próximo ao final da 1ª grande guerra mundial, ele foi amplamente utilizado para selecionar mais de um milhão de recrutas somente nos Estados Unidos, sendo naquele momento o maior sucesso que a nascente psicologia já havia produzido na forma de um instrumento científico útil e prático. Diante disto seria inequívoco que os testes psicológicos acompanhassem a formação acadêmica de tais profissionais, com o intuito de se melhorar os testes existentes e de elaborar novos testes com novos propósitos (Gardner, 1995).

Como qualquer procedimento técnico, é desejável que os testes psicológicos sejam formulados e construídos através de extensa pesquisa científica, utilizando procedimentos de padronização indispensáveis para que desta forma pudessem produzir resultados seguros. Assim qualquer teste psicológico precisa ser aprovado por avaliações de validade e precisão, para somente após esta etapa atingir o nível de teste psicométrico, ou seja, possuir a capacidade de realmente estudar aquilo a que se propõe.

Noronha et al (2003) defende que os testes psicológicos necessitam de estudos de verificação dos parâmetros psicométricos, produzindo assim normas regionais de estudos de validade e precisão, mas infelizmente no Brasil tal prática ainda não está totalmente difundida, contudo conforme estudo realizado pelos mesmos autores no que diz respeito à padronização, validade e precisão de testes psicológicos de inteligência e personalidade, constatou-se que na década de 80 e 90 no Brasil ocorreu crescimento considerável no quesito padronização dos testes psicológicos em relação a décadas anteriores, principalmente em relação aos testes de inteligência, contudo existe uma grande quantidade de testes psicológicos que ainda não passaram por estudos de validação.

Em pesquisa realizada por Castelar (1992) com profissionais que utilizam testes psicológicos, verificou-se que existem várias críticas quanto à disseminação dos testes no Brasil, uma vez que em muitos testes apenas se realizaram adaptações as quais podem contribuir para uma baixa qualidade na aplicação e no resultado dos testes, bem como na difusão das respostas utilizadas pelos próprios instrumentos, falseando assim o empenho e característica real de determinado candidato que já possua conhecimento de tais respostas. Desta forma, diversos outros instrumentos alternativos foram empregados ao longo do tempo nas mais diversas áreas da psicologia, como por exemplo na área da psicologia organizacional para seleção de pessoas, substituindo assim instrumentos padronizados e cientificamente mais concretos.

Talvez o grande empecilho do uso de testes não padronizados seja a grande chance de interferência do avaliador na aplicação ou interpretação de tal instrumento, provocando dificuldade de compreensão nos candidatos e por conseguinte descrédito científico por tal instrumento. Desta forma percebe-se que no Brasil existe a necessidade de realização de mais estudos para o desenvolvimento de novos instrumentos psicométricos, bem como levar em consideração os desvios estatísticos que existirão por conta da realidade econômica e social de cada região componente desta nação, pois ainda há necessidades de realização de pesquisas fundamentadas na padronização da coleta e análise de dados tanto de nível nacional como regional, permitindo desta forma que os pesquisadores da área de avaliação psicológica possam realizar estudos estatísticos mais profundos e precisos acerca da realidade brasileira sobre avaliação psicológica.

Segundo Baumgart (2004) percebe-se que a realidade brasileira quanto à utilização e regulamentação de testes é algo bastante recente. Somente no ano de 2003 (através da resolução 002/2003) o Conselho Federal de Psicologia regulamentou a elaboração, utilização e a

consequente comercialização de instrumentos de uso psicológico no Brasil, prevendo que tais testes devam ser avaliados por uma comissão que tem o objetivo de analisar requisitos científicos básicos para permitir sua utilização pelos profissionais da psicologia. A resolução aponta ainda a necessidade de atualização e adaptação dos testes à população que se pretende avaliar, objetivando que tais instrumentos possam medir fidedignamente os resultados produzidos pelos participantes.

Ainda segundo Castellar (1992) os instrumentos utilizados em testagem psicológica se propõem a medir diferenças ou características individuais, sendo que tais instrumentos devem possuir os atributos de fidedignidade (grau em que duas medições separadas e independentes concordam entre si ou a capacidade do teste apresentar resultados semelhantes em várias aplicações em períodos distintos no mesmo indivíduo). Já em se tratando do atributo de validade, as pesquisas com instrumentos de testagem devem fornecer evidências empíricas que sustentem as interpretações possíveis dos escores nos testes.

As pesquisas utilizando testes e/ou avaliações psicológicas podem possuir análises quantitativas ou qualitativas. Em relação aos estudos quantitativos é de conhecimento geral que a ciência moderna fundamenta-se em modelos matemáticos para explicar e avaliar os fenômenos observados e na psicologia a estatística torna-se ferramenta primordial para atender as necessidades de uma pesquisa científica, dando início a chamada psicometria como o ramo em que o modelo quantitativista se insere. No modelo psicométrico clássico temos a Teoria Clássica dos Testes (TCT) também conhecida como Teoria Clássica de Medidas (TCM), a qual se fundamenta no escore do indivíduo em um teste, sendo este escore considerado um conjunto de comportamentos (Pasquali, 2003).

Para Primi (2010), Rodgers (2010) e Muniz (2004) a área da avaliação e testagem psicológica constitui-se no ramo responsável pela formulação e aperfeiçoamento de técnicas e instrumentais, possibilitando a operacionalização de determinadas teorias psicológicas, favorecendo a observação sistemática de eventos comportamentais bem como integrando e aprimorando a construção do conhecimento psicológico teórico com o prático. A avaliação psicológica consiste num processo amplo, objetivando a busca pelo entendimento do funcionamento psíquico utilizando para isto testes e procedimentos fidedignos e metodológicos, os quais permitirão a utilização de dados ou indicadores do comportamento humano que possam ser os mais genuínos possíveis. Dessa maneira existe a evolução tanto teórica como técnica dos modelos de testagem, bem como a própria produção teórica da psicologia permite-se a discussão

de validação ou não de hipóteses exaustivamente avaliadas, aproximando desta forma a psicologia aos mais diversos ramos precisos e confiáveis da produção científica, e a Teoria Clássica de Medidas passou a ser utilizada como modelo estatístico principal por décadas.

É sabido que na área da psicologia o século XX teve a inteligência como o centro dos estudos psicológicos, sendo o nome de Charles Spearman, através de suas investigações sobre tal temática, conhecido como o desenvolvedor da metodologia da Análise Fatorial. A análise fatorial tem o objetivo de avaliar a dimensionalidade dos instrumentos de medida através de técnicas estatísticas, no intuito de descrever uma gama de variáveis iniciais através da utilização de um pequeno número de fatores (variáveis hipotéticas), os quais são obtidos a partir da função das variáveis iniciais. Assim a análise fatorial busca facilitar o entendimento existente das relações entre as variáveis através dos fatores encontrados. Diante desta nova forma de analisar dados, ocorreu crescente interesse pelo desenvolvimento de instrumentos na área da psicologia que pudesse captar construtos pertencentes a distintas dimensões cognitivas (Junker & Sijtsma, 2006; Nunnally, 1970; Sartes & Souza-Formigoni, 2013; Spearman, 1904).

Contudo, a Teoria Clássica de Medidas produziu ao longo do século XX formas longas de avaliação psicológica e até certo ponto repetitivas, o que era desejado para a época em questão, objetivando melhorar as propriedades psicométricas dos tipos de testes então vigentes, mas que se mostrou desnecessário a partir da década de 70 quando surgiram novas metodologias utilizando escalas mais curtas ao mesmo tempo em que demonstravam enorme confiança na apuração dos dados, ou seja, o tamanho das avaliações diminuiu sem que se diminuísse a sua credibilidade psicométrica (Embretson & Reise, 2000; Pasquali & Primi, 2003). A TCM também apresentava o inconveniente científico de que as próprias amostras de indivíduos participantes dos diversos tipos de modelos de testes poderiam, até certo ponto, influenciar as medidas que os testes se propunham a realizar, ou seja, as medidas seriam dependentes da amostra dos sujeitos participantes, condicionando a sua validade a uma amostra representativa da população ou ainda que o mesmo instrumento tivesse sido aplicado em uma outra amostra que possuísse características próximas à primeira. Percebe-se que tais limitações favoreceram o surgimento de pontos de desconfiança que nunca escaparam às discussões acerca dos dados encontrados em pesquisas utilizando a TCM (Pasquali & Primi, 2003; Sartes & Souza-Formigoni, 2013; Vendramini, Silva & Canale, 2004).

Por conta das limitações observadas ao longo das décadas na TCM foram propostas novas formas de abordagens psicométricas sem que fosse necessário entrar em conflito com a própria

TCM, mas apenas superar suas limitações já conhecidas. Ainda no século XX Thurstone realizou pesquisas que contribuíram com uma metodologia estatística que hoje é considerada a precursora probabilística da Teoria de Resposta ao Item (TRI), iniciando a era da psicometria moderna na década de 80. Surge então a Teoria de Resposta ao Item a qual difere da Teoria Clássica de Medidas pelo fato de analisar o sujeito a partir do seu desempenho em cada item do teste respondido, não levando em conta o resultado global no teste, buscando assim superar algumas limitações da TCM bem como proporcionando melhorias na qualidade da avaliação da estrutura dos testes (Thurstone, (1927, 1928); Spearman, (1904); Sartes & Souza-Formigoni, (2013); Araújo, Andrade e Bortolotti (2009).

A Teoria de Resposta ao Item (TRI) é considerada mais moderna que a Teoria Clássica de Medidas (TCM) pois se fundamenta no modelo do traço latente o qual é um conjunto de aptidões que o sujeito possui e que auxilia no seu desempenho ao responder uma tarefa do teste em questão (Pasquali, 2003). Este novo modelo de análise vem galgando espaço dentro da análise de dados nos testes e avaliações psicológicas, fomentando a construção de escalas de medidas confiáveis e auxiliando os pesquisadores a entender melhor ou até mesmo a reformular hipóteses científicas. A TRI se propõe a analisar estatisticamente cada item de forma particular, tentando minimizar a dependência exclusiva do teste, mantendo a importante característica de não entrar em conflito com princípios que formam a Teoria Clássica de Medidas, mas apenas focando-se em obter escalas sem a necessidade de utilização de um escore geral obtido num teste.

Segundo Fernández (1990) e Muñiz e Hambleton (1992), a TRI auxilia a elevar a qualidade métrica dos instrumentos de avaliação, bem como adequar seu rigor levando-se em conta as especificidades dos participantes da pesquisa através da organização dos itens do teste a partir de duas funções matemáticas: a logística e a normal padronizada, as quais fornecem dados sobre os parâmetros dos itens a partir de curvas características.

Segundo Fletcher, (1994), Embretson e Reise, (2000) a TRI fundamenta-se na existência do traço latente no sujeito, chamado também de construto, ou seja, diz respeito a uma característica presente neste indivíduo e que possui níveis, os quais serão aferidos em termos probabilísticos nas respostas do indivíduo a cada item da avaliação. Na TCM o cálculo do escore do indivíduo sofria variações dependendo da precisão e do nível de dificuldade do instrumento utilizado, enquanto a TRI considera que existe um traço latente o qual é fidedigno e específico, ou seja, não estará sujeito às variações observadas em testes fundamentadas pela TCM.

Para Andriola, (1998) é preciso levar em consideração dois aspectos característicos da TRI: o primeiro fala a respeito de que o grupo de itens presente no teste pretende realizar a medição de um mesmo traço latente utilizando a análise fatorial, enquanto um outro aspecto importante fala sobre a independência local dos itens, ou seja, a resposta que o sujeito apresentou a um item não vem influenciada ou possui capacidade de influenciar a resposta de outros itens do teste.

Sabe-se que o ser humano é um ente comportamental multideterminado e assim a idéia da existência de um único traço latente que seria o responsável pelo comportamento do sujeito no teste torna-se pouco satisfatória, contudo os modelos na TRI que utilizam o postulado da unidimensionalidade falam a favor da existência de um fator ou traço latente que seria dominante em relação aos demais traços, mas não o único. Desta maneira credita-se a essa dominância de um dado traço latente as respostas comportamentais do sujeito ao teste aplicado, porém ainda persistem opiniões que discutem o que seria de fato a unidimensionalidade e como ela poderia ser melhor analisada e assim surgiram modelos multidimensionais que se propõem a avaliar ao mesmo tempo múltiplas dimensões do comportamento humano, agilizando, facilitando e evitando para os pesquisadores a necessidade de aplicação de vários testes que teriam como objetivo a coleta de apenas uma única dimensão do comportamento Pasquali e Primi, (2003); Andrade, Tavares e Valle, (2000).

Cabe ressaltar que os modelos matemáticos produzidos a partir da TRI para os traços latentes visam representar a possível relação probabilística da resposta dada a um item com o seu traço latente e parâmetros dos itens, permitindo assim uma escala de medida. Uma vantagem da TRI sobre a TCM consiste na possibilidade de realizar comparações entre traço latente de indivíduos de populações distintas ou mesmo o comparativo de indivíduos de uma mesma população que se submeteram a testes distintos, uma vez que a TRI se fundamenta em itens e não no teste como um todo, permitindo também que o pesquisador conheça os itens que estão formando a informação contida na escala, bem como os dados que não foram respondidos podem passar por um tratamento a partir das informações dos itens que foram respondidos. Dentre os modelos existentes de TRI pode-se classificá-los a partir dos tipos de processo de resposta, onde encontraremos o tipo de TRI acumulativo (onde a probabilidade do participante fornecer ou escolher uma resposta correta ao item pesquisado aumenta com o aumento do traço latente) e o tipo de TRI desdobramento (baseados em modelos de resposta onde o participante tende a

selecionar a opção de resposta que é a mais próxima da sua posição do traço latente) Vendramini, Silva e Canale, (2004), Araújo, Andrade e Bortolotti, (2009).

O primeiro tipo foi desenvolvido buscando sua utilização principalmente em medidas para avaliação na área educacional, onde se observa que o traço latente mais pesquisado por este tipo de pesquisa é a habilidade ou proficiência, a qual possui probabilidade de aumento no indivíduo no transcorrer da avaliação que está sendo aplicada. Já a distinção existente entre este tipo e a TRI modelo de desdobramento reside na característica de que esta última foi projetada visando a análise de dados de atitude, bem como se propõe a aproximar as respostas do indivíduo ao seu provável traço latente real, ou seja, existe uma medida probabilística entre os parâmetros do indivíduo em relação à posição do item na escala em questão e desta forma deve existir uma pontuação individual e inerente ao indivíduo na escala medida para um dado traço latente, onde a resposta deste indivíduo deverá se localizar a pouca distância do ponto ideal do indivíduo, buscando desta forma fugir daquilo que seria o parâmetro deste indivíduo, tal como acontece na TRI de modelo acumulativo (Andrade & Bortolotti, 2007).

Em suma, o modelo de análise a partir da TRI auxiliou a busca por avanços nas medidas de traços latentes, uma vez que possui características vantajosas em relação à Teoria Clássica de Medidas, pois permitiu aos pesquisadores caracterizar a natureza do traço latente que se deseja analisar bem como construir itens que possam abranger o traço latente e interpretar sua escala surgida a partir dos próprios itens. Também pode-se apontar que a TRI possibilitou a otimização de instrumentos que já haviam sido validados a partir da TCM, ou seja, a TRI pode complementar e auxiliar a TCM, sobretudo os processos de análise. Todavia, a TRI é considerado um método pouco simples quando comparado à TCM, uma vez que exige compreensão por parte do pesquisador para que a análise de interpretação dos dados entre os itens e o traço latente (que se espera encontrar) possam produzir informações psicométricas sólidas. Segundo Sartes et al (2013) a TCM e a TRI podem avaliar de maneira profunda e se mostrar muito eficientes quando aplicadas em conjunto, possibilitando a formação de dados mais abrangentes e minuciosos em distintas etapas de análise de pesquisa (Araújo, Andrade & Bortolotti, 2009; Sartes & Souza-Formigoni, 2013).

Dito isto, observa-se a importância da área de avaliação psicológica para o aperfeiçoamento teórico e instrumental da psicologia, valorizando as teorias validadas e possibilitando profundas discussões dos pontos não-validados pela análise matemática aplicada aos

testes, mas de toda forma permitindo a construção científica dentro da própria psicologia (Junker & Sijtsma, 2006; Sartes & Souza-Formigoni, 2013; Spearman, 1904).

Assim observa-se que o contínuo desenvolvimento estatístico e psicométrico dos testes e de seus modelos de análise tornou-se imprescindível para que o psicólogo dispusesse de instrumentos com boas propriedades psicométricas e que pudessem ser usados nas avaliações dos mais diversos construtos, auxiliando na fundamentação de tomadas de decisões profissionais. Dessa forma esse trabalho se propõe a realizar uma análise psicométrica de um novo instrumento para avaliação do Raciocínio Lógico-dedutivo (um aspecto do Pensamento Crítico), tanto com base na TCT quanto na TRI.

A psicologia enquanto ciência tem o homem como objeto de estudo em seus aspectos cognitivo, emocional e comportamental e, no século XX, a busca pelo entendimento dos processos cognitivos tornou-se o cerne dos estudos psicológicos. Dentre esses processos, a inteligência vem sendo amplamente estudada pelos mais diversos pesquisadores de diferentes países, possibilitando não somente seu melhor entendimento como também o desenvolvimento de instrumentos capazes de medi-la com maior eficácia, especialmente as formas de inteligência conhecidas como linguística (verbal) e a lógico-matemática (fluída). Isto possibilitou o desenvolvimento de uma formação curricular educacional fundamentada nesses tipos de inteligência, bem como a elaboração e desenvolvimento de testes psicológicos que visavam medir e acompanhar o desenvolvimento deste processo cognitivo (Bock, Furtado & Teixeira, 1999).

A inteligência é um construto biopsicossocial, sendo em última análise uma produção de ideias pessoais reelaboradas e repensadas a partir da ótica de movimentos filosóficos, culturais e científicos de determinada época. Assim, pode-se atualmente valorizar certos aspectos do conceito de inteligência que em outros séculos não seriam valorizados ou até mesmo percebidos como tal. Pode-se, então, pensar e concluir que qualquer instrumento que visa medir a inteligência humana não o faz em absoluto, pois na verdade o que se está medindo não seria a inteligência pura, límpida e direta, mas sim os conceitos vigentes sobre a inteligência que estejam sendo valorizados por determinado grupamento humano, ou seja, avalia-se apenas um constructo a partir de alguma perspectiva teórica. Ainda assim, existem correlações matemáticas que qualquer teste poderá se utilizar, galgando maior credibilidade aos seus achados e possibilitando diversos tipos de comparações com outras experiências científicas realizadas por outros pesquisadores em diversas regiões do planeta e desta forma um novo corpo de conhecimentos vai sendo adicionado à este

construto que se mostra em indefinida e lenta mutação, facilitando o entendimento do homem sobre aspectos individuais e gerais do funcionamento da mente humana (Balbinotti, 2005).

Para Gardner (1998, p.215) a inteligência se define como “a capacidade de resolver problemas ou criar produtos que são importantes num determinado ambiente cultural ou comunidade” e por isso não deve existir um único tipo de inteligência, mas sim inteligências múltiplas, as quais seriam: a inteligência linguística, a inteligência musical, a inteligência lógico-matemática, a inteligência espacial, a inteligência corporal-cinestésica, inteligência intrapessoal e a inteligência interpessoal, cada uma com especificidades e formas de evidência próprias. Assim um teste que verificasse o potencial de inteligência lógico-matemática de um indivíduo pode não ser indicado para se verificar o potencial da inteligência interpessoal no mesmo indivíduo, pois a manifestação destas inteligências se dá de formas distintas, embora as duas sejam importantes para a sobrevivência do indivíduo. Os testes que medem o quociente de inteligência (QI) foram os que mais se aprimoraram ao longo do século XX, tornando-se instrumentos científicos úteis e práticos (Gardner, 1995).

Esses instrumentos vêm sendo amplamente investigados há décadas, sendo fundamentais para a própria constituição da área de avaliação psicológica (Primi, 2002). No entanto, há um campo ainda pouco explorado, especialmente no Brasil, relacionado com o raciocínio em situações cotidianas, denominado de pensamento crítico, que se baseia numa estrutura da lógica informal. Diversas pesquisas relacionadas a esse tipo de processamento cognitivo da lógica informal se relacionam com a argumentação e o pensamento reflexivo e tem sido desenvolvida em diversos centros acadêmicos (inclusive na Universidade Federal de Pernambuco no âmbito do Núcleo de Pesquisas em Argumentação - NuPArg), utilizando metodologia de pesquisa qualitativa, analisando também as perspectivas no estudo da argumentação quotidiana e o papel do professor na construção discursiva da argumentação em sala de aula (De Chiaro & Leitão, 1995; Leitão, 1996, 2007). Apesar das importantes contribuições que essas pesquisas têm dado à área de estudos do raciocínio informal, elas se referem mais especificamente à aplicação da argumentação em sala de aula. Diante disto foi percebida a existência de carência de pesquisas que utilizassem metodologia quantitativa acerca da lógica informal focando o pensamento crítico. Um exemplo dessa carência é a existência de poucos estudos que avaliem as propriedades psicométricas para avaliação desse construto no contexto cultural brasileiro e, por isso, esta pesquisa procura contribuir com produção acadêmica neste tipo de avaliação.

O pensamento crítico e seus conceitos

Segundo Kopnin (1978) o fazer humano transforma a realidade objetiva e a razão atuaria como o principal fator direcionador dessas ações. Por isso a razão se definiria como forma profunda do conhecimento teórico do homem sobre a realidade que o cerca, sintetizando o conhecimento do aspecto racional e estudando a própria natureza deste conhecimento. Por tal fato, se distingue do juízo, o qual constitui um dos aspectos racionais, mas que atua com a função de avaliar e desmembrar as relações objetivas que o homem consegue captar de tal realidade (Bernardes, 2011).

Diante de tais afirmações pode-se definir a razão como a faculdade tipicamente humana de argumentar, retirando de tal argumentação conclusões que podem ser verdadeiras ou falsas. Os argumentos permitem que se extraiam de pressupostos verdadeiros conclusões necessariamente verdadeiras (argumentos dedutivos) ou provavelmente verdadeiras (argumentos indutivos). Entende-se como bons argumentos aqueles em que existe a garantia de sua veracidade, eliminando assim a probabilidade. A argumentação dedutiva envolve regras que, sendo rigorosamente seguidas, tornam a argumentação irrefutável, partindo-se do princípio do geral para o particular (método dedutivo) (Nolt & Rohatyn, 1991).

Nesse âmbito a lógica informal possui como objeto o estudo de argumentos que estão dispostos em linguagem comum, diferenciando assim das apresentações de argumentos numa linguagem formal (sentenças). A lógica informal serve então para analisar ou construir formas de argumentação que são comumente utilizadas na rotina da vida diária, permitindo entre os comunicantes a validação ou não da mensagem, ou seja, discernir o argumento válido das falácias (argumentos não-válidos). Pode-se encontrar exemplos da lógica informal na análise de publicações diárias como jornais, livros e revistas, cujos textos breves em grande parte utilizam-se de falácias. De toda forma, o ramo da lógica informal permite diversos tipos de estudos sobre a forma com que se processa o raciocínio e é objeto de grande interesse de pesquisa, principalmente por parte das ciências humanas, uma vez que também pode auxiliar o pesquisador no entendimento de formas de raciocinar criticamente sobre os argumentos (Nolt & Rohatyn, 1991).

Dentro do campo da lógica informal existe a área de estudo do pensamento crítico, cuja expressão se refere a uma tradição de investigação de origem filosófica e de desenvolvimento multidisciplinar. Esta área foi ampliada na década de 60 quando alguns filósofos planejaram estudar várias formas possíveis de raciocínio, se preocupando não somente com as modalidades

formais, mas também com as informais, sobretudo os raciocínios de cunho cotidiano, que possuem natureza mais indutiva (que partem da análise do particular ou específico até se chegar a uma conclusão capaz de abranger o todo) do que dedutiva (chegada a uma conclusão acerca de um evento específico a partir de uma análise do todo) (Ennis, 1996).

Em suma, constata-se que o estudo do pensamento crítico se desenvolveu baseado na lógica informal, a qual se preocupa com todas as formas de raciocínio, contemplando aquelas manifestações reflexivas deixadas de lado pela lógica formal, mas que são igualmente importantes na construção do raciocínio. Dessa forma a lógica informal, ao longo dos anos, passou a ser conhecida pelo termo raciocínio informal ou raciocínio prático, sendo motivo de investigação por parte de filósofos, professores, educadores e psicólogos, sendo estes últimos os maiores responsáveis pela consolidação deste campo do saber. Vale salientar que as capacidades fundamentais do raciocínio são propostas como habilidades básicas de comprovação de hipóteses, resolução de problemas e tomada de decisão, sendo os psicólogos o grupo de pesquisadores interessados na identificação da atividade intelectual como forma de estudar o pensamento crítico propriamente dito (Saiz, 2002).

Bloom (1956) desenvolveu uma classificação das atividades cognitivas em seis tipos qualitativos: conhecimento, compreensão, aplicação, análise, síntese e avaliação. Tal classificação permitiu a realização de pesquisas no campo educacional, na qual o pensamento crítico foi incluído, permitindo a discriminação de informações relevantes no contexto do desenvolvimento educacional.

Brookfield (1987) acrescenta que o pensamento crítico não se constitui somente num ato cognitivo, mas também num ato emocional, uma vez que ansiedade, confusão, medo e resistência podem atuar durante o processo racional e, assim, o bom controle e manejo da situação emocional consegue facilitar o desencadeamento do pensamento crítico.

Por sua vez Minnich (1990) considera que o conceito de pensamento crítico talvez tenha se tornado “mistificado” (um conceito ou ideia tão habitual que raramente provoca questionamentos), o que seria contrário ao pensamento crítico, que tenta provocar questionamentos em conceitos ou ideias habituais. A autora dá como exemplo a existência de diversos conceitos acadêmicos que normalmente são utilizados sem que necessariamente se possua entendimento comum ou prático de seu mecanismo de funcionamento. No entanto, através do pensamento crítico

o indivíduo conseguiria captar mais informações acerca de tal conceito ou objeto de estudo, inclusive percebendo aquilo que na verdade não consegue entender a fundo.

McPeck (1981, p. 8) conceitua o pensamento crítico como “uma tendência ou habilidade de se engajar numa atividade por ceticismo reflexivo/refletido”. Tal definição enfoca atitudes e habilidades próprias do pensamento crítico, uma vez que o próprio pensar criticamente sobre algo surge como produto de reflexão profunda acerca deste mesmo algo. Ennis (1989, p.4) enfatizou as ações de um sujeito que pensa, definindo o pensamento crítico como “um pensamento refletido moderado focado na decisão do que acreditar ou de como agir”.

Halonen (1995, p. 76) propôs uma definição que unifica os conceitos de McPeck (1981) e Ennis (1989), na qual o pensamento crítico seria uma “tendência ou habilidade de se engajar numa atividade por ceticismo reflexivo/refletido focado na decisão do que acreditar e de como agir”.

Halpern (2003) propõe que o pensamento crítico possua a atividade de comportamentos ou condutas que possibilitem o alcance de um objetivo específico. Essas condutas passam pela compreensão de uma linguagem e análise de argumentos que foram dados, possibilitando assim testar hipóteses, probabilidades e incertezas, culminando com o momento da tomada de decisão, visando posterior resolução de uma situação problemática que desencadeou todo o processo. Nesse sentido, Paul (2005) enfoca o final do processo, definindo o pensamento crítico como um processo de tomada de decisão que é capaz de produzir boas decisões.

Brady (2008) formulou o conceito de que o pensamento crítico se caracteriza por um tipo de raciocínio de complexidade e qualidade superiores, capaz de articular conhecimento, experiência e competências intelectuais. Essa complexidade, segundo Barnes (2005), faz com que o pensamento crítico se torne particularmente relevante no processamento dos conteúdos exigidos no ensino superior.

Em relação aos conceitos sobre pensamento crítico anteriormente relatados, nota-se que esta habilidade cognitiva possui sua importância de desenvolvimento dentro do contexto do sistema educacional, ou seja, durante o percurso de qualquer indivíduo na escola desde a mais tenra idade, percebe-se que as alterações comportamentais (sobretudo aquelas que envolvem a aquisição de conhecimentos), passa por processos e mudanças profundas no que diz respeito à didática e à pedagogia, com novas exigências sendo propostas no decorrer das faixas etárias e consequentemente para cada nível escolar. Entende-se que em se tratando da educação do nível

superior, o estudante precisa desenvolver habilidades que possibilitem o enfrentamento de novas situações em seu futuro cotidiano profissional, bem como desenvolver a capacidade de propor resoluções adequadas para cada tipo de problema que poderá ser vivenciado, o que também demanda aspectos comportamentais que falam a respeito de uma participação ativa no que tange à sua experiência tanto acadêmica quanto no próprio ambiente de trabalho.

Assim, o indivíduo buscaria atingir autonomia e sobretudo iniciativa (ainda enquanto estudante) acerca não apenas dos possíveis problemas que necessitarão de resolução, mas também dos aspectos referentes ao seu processo de aprendizagem durante seu período de vida universitária (Carelli & Santos, 1998). Pesquisas realizadas por Carelli e Santos (1998) e por Vasconcelos (2005) sugerem que uma enorme gama de estudantes que ingressam no nível superior de ensino no Brasil apresentam dificuldades de aprendizagem e que tais dificuldades podem estar relacionadas aos métodos utilizados por este nível acadêmico, dificultando assim a adaptação deste aluno a esta nova realidade de ciclo escolar.

É possível então admitir que tais alunos têm possibilidade de prejuízo quanto ao seu potencial rendimento global de aquisição e desenvolvimento deste conhecimento que será específico para sua atuação profissional. Sabe-se que o nível do ensino superior exige não apenas maior domínio de temas e assuntos já abordados nos outros ciclos escolares, mas principalmente novos métodos de estudo elaborados pelo próprio aluno, o que implica numa necessidade de autonomia cada vez maior no avançar de etapas deste ciclo, a qual por sua vez depende da capacidade de autorregulação de cada indivíduo e que a evolução acadêmica para os níveis de pós-graduação também demandará capacidades cada vez mais eficazes de autorregulação do estudante, ou seja, o próprio aluno necessita apreender conteúdos cada vez mais complexos e de maneira mais rápida e autônoma, sendo tal habilidade imprescindível neste ciclo acadêmico (Almeida; Guisande; Pereira; Joly; Donaciano; Mendes; Ribeiro, 2009).

Segundo Zimmerman (2000) a autorregulação constitui qualquer sentimento, ação ou pensamento que pode ser orientada pela própria pessoa objetivando o alcance de seus objetivos. No caso de estudantes, o objetivo seria aprender determinado conteúdo a partir da seleção apropriada de estratégias eficazes e adequadas às suas necessidades singulares, bem como de sua realidade social, com conseqüente revisão organizada e sistematizada destas mesmas estratégias, possibilitando novos direcionamentos para o alcance do objetivo proposto: aprender, assimilar e refletir acerca de novos e complexos conteúdos inerentes à sua profissionalização.

Meece (1994) e Winne (1995) apontam que o controle sobre o afeto e a conduta ativa acerca da aquisição do conhecimento, bem como o controle do próprio pensamento, com capacidade de escolha de estratégias para atingir objetivos propostos, acompanhada da tomada de consciência pelo indivíduo acerca das dificuldades que podem vir a impedir o seu processo de aprendizagem (ou seja, suas próprias limitações pessoais de cunho cognitivo, social, afetiva, etc), constituem componentes imprescindíveis ao conceito de aprendizagem autorregulada. Considera-se autorregulados aqueles indivíduos que conseguem organizar, implementar e controlar um planejamento de sua metodologia específica de estudos, nunca abandonando o ideal principal que é o alcance de seus objetivos e metas iniciais, sendo conseqüentemente importante que o mesmo indivíduo utilize também a autoavaliação acerca deste método de estudo o qual ele próprio selecionou e está vivenciando.

A importância educacional relacionada ao pensamento crítico reside no fato de que este tipo de raciocínio possui a capacidade de potencializar o processo de tomada de decisão, permitindo que o problema em questão tratado pelo indivíduo seja conduzido através da resolução com maior eficácia, o que resulta em aspectos comportamentais e cognitivos desejáveis no contexto de aprendizagem escolar, pois permitem monitoramento do desempenho do aluno bem como avaliar sua capacidade de auto-análise, ou seja, permite que o próprio aluno consiga, paulatinamente e desde cedo, desenvolver sua autorregulação, mesmo que esta habilidade seja mais valorizada no ensino superior, nada impede que sua importância possa ser atrelada ao pensamento crítico em qualquer ano escolar, uma vez que o pensamento crítico poderá ser desenvolvido independentemente da série em que o aluno se encontra, porém a metodologia educacional considerada mais dependente da autorregulação (e conseqüentemente do pensamento crítico) ainda é a metodologia de ensino utilizada no ensino superior, sendo os alunos deste nível acadêmico exaustivamente treinados com foco em desenvolvê-la (Lindner & Harris, 1993; Barnes, 2005; Saiz & Rivas, 2010; Pintrich, Smith, Garcia & McKeachie, 1993; Paul, 2005).

O pensamento crítico e seus fatores constituintes

Halonen (1995) chama atenção para o fato de que existem três tipos de abordagens ou escolas que utilizam o pensamento crítico com perspectivas distintas: a básica, a aplicada e a crítica social. A Escola Básica adota uma perspectiva da psicologia educacional, explicando o pensamento crítico através de um viés predominantemente cognitivo. A Escola Aplicada se utiliza de uma perspectiva pedagógica, que foca no aperfeiçoamento tanto do ensinar como do aprender. Por sua vez, a Crítica Social se coloca como proponente de uma reforma educacional, social e política incluindo o pensamento crítico, ou seja, não o concebe como algo sem contextualização.

Essas escolas assentam-se, respectivamente, em três formas de se explorar a essência do pensamento crítico: traço, emergente e condição/estado. A primeira está relacionada com as linhas tradicionais de pesquisas na área da inteligência, identificando diferenças individuais nas inclinações naturais ao pensamento crítico. A forma emergente deriva das pesquisas de Piaget bem como de outras pesquisas na área de desenvolvimento, defendendo que a habilidade do pensamento crítico emerge da interação da criança com o ambiente, desdobrando-se naturalmente através do processo de desequilíbrio. Já a linha de condição/estado considera o pensamento crítico como atitudes discretas do comportamento e enfatizam a demonstração do pensamento ao invés do seu desenvolvimento (Hallonen, 1995).

Mentkowsky (1991) afirma que tais linhas ou formas de explorar a essência do pensamento crítico têm em comum o fato de concordar que este se constitui numa capacidade multidimensional e não numa simples habilidade isolada. Em pesquisas mais recentes o pensamento crítico é entendido como mediador da tomada de decisão em diversos tipos de situações da vida cotidiana, auxiliando na identificação da melhor solução para o problema enfrentado (Saiz & Rivas, 2010). Carroll (2005) salienta que a utilização apropriada do pensamento crítico possibilita ao indivíduo definir um objetivo que se pretende atingir através da utilização de recursos disponíveis, como o próprio conhecimento prévio acerca do tema em questão, desenvolvendo assim raciocínio claro e fundamentado que possibilite seu movimento na direção da resolução mais adequada ao que está sendo exigido.

Para West, Toplak e Stanovich (2008) o pensamento crítico é enquadrado como um subdomínio do pensamento racional, consistindo numa aptidão que serve de interface em processos

de tomada de decisão para resolução de problemas cotidianos. O pensamento racional, tal como proposto pelos autores, engloba uma linha epistêmica associada à construção e à capacidade de alcançar objetivos bem como ao desenvolvimento de estratégias adequadas para a resolução de problemas. Nesse sentido, o pensamento racional é considerado como um processo que pretende explicar a razão das tomadas de decisão, fundamentadas na experiência do sujeito. Esses autores chamam a atenção ao fato de os raciocínios heurísticos constituírem domínios cognitivos que integram aspectos processuais de raciocínio causal, hipotético, científico, disjuntivo, pensamento argumentativo, pensamento hipotético e capacidade de procurar explicações alternativas à resolução de problemas. Contudo, tais domínios não são avaliados conjuntamente aos testes de inteligência e de pensamento crítico, mas os autores acreditam que tais processos cognitivos possuem implicação nesta forma de pensamento.

Rivas e Saiz (2012) consideram que habilidades cognitivas capazes de discriminar a informação relevante da não relevante, bem como a capacidade de elaboração e valoração argumentativa e também o correto julgamento de eventos probabilísticos bem como a capacidade de perceber correlações entre estes eventos constituem-se em importantes atributos que auxiliam o indivíduo a resolver questões de natureza complexa como também de natureza cotidiana, e seriam agrupadas e estudadas dentro de um conceito denominado pensamento crítico. O pensamento crítico não seria então uma forma de pensar automatizada, mas sobretudo reflexivo e propositivo. Estes autores concebem o pensamento crítico como uma teoria da ação, uma vez que pensar criticamente implica em resolver problemas e alcançar metas, não ficando somente no campo do bom juízo ou da argumentação, sendo este um meio e não um fim para o alcance de uma ação. Assim, a resolução de problemas com eficiência está alicerçada na tomada de decisão bem fundamentada, a qual demanda reflexão profunda. Esta conceituação desenvolvida por Rivas e Saiz será adotada para a realização desta pesquisa, uma vez que a mesma foi considerada como a mais completa e a mais atualizada dentre as conceituações sobre o pensamento crítico já apresentadas neste trabalho.

Os estudos acerca das habilidades gerais do pensamento de ordem superior (o pensamento crítico e criativo) evoluíram de maneira menos acelerada em comparação com o rápido desenvolvimento observado nos procedimentos de investigação das demais estruturas cognitivas, principalmente quando comparado aos estudos sobre a inteligência (Greeno, 1989).

A partir do conteúdo exposto, pode-se perceber que o pensamento crítico se insere em uma área ainda emergente dos estudos da cognição, no entanto o reconhecimento crescente de sua importância, especialmente no que se refere aos processos de raciocínio informal, se dá exatamente pelo impacto que causa em aspectos importantes da vida cotidiana, como por exemplo, nas tomadas de decisão e de aprendizagem.

A avaliação do Pensamento Crítico

O surgimento dos testes psicológicos permitiu que a psicologia estabelecesse um referencial objetivo que pudesse amenizar as percepções subjetivas do avaliador e que também conseguissem ter seu uso disponibilizado para diferentes contextos da pesquisa psicológica, proporcionando resultados céleres e o mais objetivos possíveis. A busca por um modelo eficiente de medição de traços psicológicos (propriedades psicológicas individuais) constitui um amplo interesse compartilhado por diversas áreas de conhecimento, as quais desenvolveram técnicas de medição diversas e até certo ponto confiáveis no que diz respeito à sua eficácia estatístico-científica (Sartes & Souza-Formigoni, 2013; Araújo et al, 2009).

Em pesquisa científica que utilize análise estatística sempre teremos três critérios primordiais que deverão fundamentar a capacidade de testagem da ferramenta que foi empregada no estudo, critérios estes que são a validade, a fidedignidade e a padronização. A validade fala a respeito da relação que ocorre entre o estudo (teste) e o que ele propõe avaliar, podendo ser uma validade de conteúdo, de critério ou ainda uma validade de constructo. O princípio da fidedignidade ou precisão pode ser descrito como o estudo da metodologia mais precisa possível, ou seja, esse princípio nos fala a respeito da confiabilidade dos resultados obtidos na testagem, permitindo assim nos mostrar consistência e quantidade mínima de erros de mensuração. Por fim, teremos o critério de padronização o qual é o procedimento que irá estabelecer a aplicação e a avaliação do teste de maneira uniforme em todos os indivíduos que serão testados, ou seja, permite assim a correta aplicação da ferramenta escolhida bem como a possibilidade de comparações posteriores por pesquisadores que utilizaram o mesmo teste em localidades e populações distintas (Urbina, 2007).

As principais dificuldades na avaliação do pensamento crítico dizem respeito a aspectos tanto conceituais como metodológicos, uma vez que os diversos conceitos sobre o pensamento crítico podem levar a diferentes formas de avaliação desta habilidade. Muitas dessas formas de avaliação, no entanto, são baseadas em instrumentos com respostas fechadas (marcação de um X), que podem impedir a exploração de aspectos importantes do processamento necessário para a resolução dos problemas apresentados (Ennis, 1987, 2003).

Para Yanchar, Slife e Warne (2008) a abordagem geral do pensamento crítico, incluindo sua avaliação, deve abranger o diálogo crítico e o confronto de divergências e pontos de vista, permitindo assim a sua abordagem de forma mais ampla e equilibrada.

O *Halpern Critical Thinking Assessment Using Everyday Situations* – HCTAES (Halpern, 2006) é um instrumento que se propõe a lidar com a questão metodológica da avaliação do pensamento crítico por meio da apresentação de problemas corriqueiros, que exigem tanto respostas abertas quanto fechadas do participante. Este instrumento foi desenvolvido considerando que o pensamento crítico seja composto por cinco habilidades: comprovação de hipóteses, raciocínio verbal, análise do argumento, probabilidade e incerteza e tomada de decisão. Este teste avalia a qualidade do pensamento crítico que subjaz a cada decisão tomada pelo participante em sua análise das situações cotidianas apresentadas (Halpern, 1998, 2010).

Em pesquisa realizada por Nieto, Saiz e Orgaz (2009) os autores contaram com uma amostra composta por 335 participantes universitários e encontraram confiabilidade medida pelo Alfa de Cronbach nos seguintes valores: 0,635 para a habilidade de comprovação de hipóteses, 0,367 para a habilidade de raciocínio verbal, 0,465 para a análise do argumento, 0,340 para probabilidade e incerteza e por fim o valor de 0,487 para a habilidade de tomada de decisão. Tais índices encontrados não estão adequados com os valores mínimos exigidos por critério internacional para um instrumento (Urbina, 2007). Este estudo também encontrou medida de adequação da amostra através da prova de Kaiser-Mayer-Olkin com valor de 0,70 e prova de esfericidade de Bartlett ($\chi^2=2214,44$, $p<0,05$). Após análise dos componentes principais os pesquisadores concluíram que não havia unidimensionalidade, ou seja, o pensamento crítico seria composto por um conjunto de habilidades e encontraram a mesma quantidade de fatores obtidas por Halpern (2006), que foram cinco. Porém os autores salientaram que a estrutura fatorial obtida nesta pesquisa distinguiu-se da estrutura fatorial teórica de Halpern, pois na formação destes 5 fatores não ocorreu a mesma distribuição de itens por fator como proposto por Halpern. Diante

disto uma outra conclusão apontada pelos autores foi a de que o pensamento crítico não se constitui numa única habilidade, mas funciona como uma, a partir de uma integração entre estas cinco habilidades de processamento cognitivo.

O *Watson-Glaser Critical Thinking Assessment* (WGCTA) constitui-se numa bateria de 80 itens de resposta múltipla, possuindo cinco dimensões avaliativas: inferências, reconhecimento de pressupostos, dedução, interpretação e avaliação de argumentos. Esta bateria apresenta também raciocínios silogísticos no qual é solicitado ao participante apontar a alternativa correta. Em estudo realizado com 793 alunos por West, Toplak e Stanovich (2008) foram encontradas desde baixas a moderadas correlações entre as cinco dimensões avaliadas pelo WGCTA e o *score* global de rendimento acadêmico, o que foi interpretado pelos autores da pesquisa como um fato ocorrido graças à pouca estimulação dada ao pensamento crítico durante o trajeto acadêmico de tais alunos.

Em outra pesquisa realizada por Bernard et al (2008) correlacionando as cinco dimensões avaliativas do WGCTA os pesquisadores encontraram significativas correlações médias, mas todas se apresentaram com pouca heterogeneidade, o que estimulou os pesquisadores a realizarem análises de componentes principais em outras medidas de subescalas no teste. Por fim os autores concluíram que o *Watson-Glaser Critical Thinking Assessment* deve ser utilizado com cautela, pois os resultados apontaram que o teste consegue avaliar uma medida de competência geral, devendo-se evitar assim interpretações individuais das dimensões que o teste pretende avaliar.

Com base em sua concepção do pensamento crítico como uma teoria da ação, Rivas e Saiz (2012) propuseram um instrumento para avaliação do pensamento crítico, que se fundamentou em três habilidades principais: o raciocínio, a resolução de problemas e a tomada de decisão. Os autores estiveram sempre preocupados com o intercâmbio dos itens do instrumento com a realidade objetiva, uma vez que eles propunham avaliar a utilização do pensamento para resolver problemas do cotidiano da forma mais eficiente. Com base nisso, os autores desenvolveram o teste PENCRISAL, que apresenta algumas distinções em relação aos instrumentos anteriormente descritos, como por exemplo a utilização de itens que abordam situações cotidianas, a utilização de diferentes domínios envolvidos (sendo os principais o raciocínio indutivo, o raciocínio dedutivo, o raciocínio prático, a tomada de decisão e a resolução de problemas), com intenção de valorizar o grau de generalização de habilidades, o formato de respostas abertas, possibilitando a exploração de diversos processos de pensamento e, por fim, o emprego de situações-problema que possuem resposta única, permitindo avaliar o mecanismo de pensamento correspondente.

Este teste é constituído por 35 situações-problema com produção de resposta aberta, sem necessidade de utilização de termos técnicos na elaboração de respostas. Os itens representam cinco fatores relacionados ao pensamento crítico, que são o raciocínio dedutivo, o raciocínio prático, o raciocínio indutivo, a tomada de decisão e a resolução de problemas e são distribuídos aleatoriamente ao longo do teste na quantidade de sete itens por fator. A concepção de raciocínio crítico proposta por Rivas e Saiz (2012) possui relação com o raciocinar sobre problemas para resolvê-los, sendo assim, atividades inter-relacionadas e a aplicação do teste podem ser realizadas com formulário tradicional, no qual o participante escreve suas respostas de próprio punho ou ainda através de formulário eletrônico no qual o participante digita suas respostas utilizando um computador, podendo ser respondido inclusive via internet. Essa opção oferece vantagens como avaliação de diversos indivíduos simultaneamente; menor tempo utilizado para a organização dos dados colhidos e a possibilidade da realização do teste com a mesma pessoa em várias sessões, reduzindo assim os efeitos do cansaço, como a diminuição do rendimento do participante principalmente nos itens finais do teste, o que poderia prejudicar a análise estatística como um todo.

Numa pesquisa realizada por Rivas e Saiz (2012) para validação da versão espanhola do PENCRISAL, foram utilizados 715 participantes adultos, de nacionalidade espanhola, de ambos os sexos, com nível universitário e idades entre 18 e 53 anos. A consistência interna dos 35 itens (Alfa de Cronbach) obtida foi de 0,632 com $p < 0,001$, evidenciando nível aceitável à complexidade do constructo em estudo. Já o teste de precisão entre avaliadores obteve alto índice de concordância, com valores de Kappa dentro do intervalo de 0,6 a 0,9 ($p < 0,001$). A validade do construto foi investigada através de análise fatorial de maneira independente em cada um dos fatores estudados (dedução, indução, raciocínio prático, tomada de decisões e resolução de problemas) obtendo-se soluções unifatoriais em todos os casos (Rivas & Saiz, 2012). Em suma, o PENCRISAL possui o objetivo de avaliar a competência lógico-argumentativa das pessoas através de situações-problema descritas numa linguagem da lógica informal, para as quais o respondente necessita escrever uma solução baseada no contexto de tal problema e justificar sua possível solução. Este teste traz uma nova abordagem sobre estudos da cognição, possuindo critérios psicométricos bem estabelecidos quando aplicado em voluntários da Espanha e se encontra em processo de adaptação e validação em outros países, inclusive na América Latina.

Por sua vez no Brasil existe o teste de raciocínio lógico-dedutivo (desenvolvido por Primi e Muller, não publicado) o qual foi construído a partir de pressupostos lógicos que possibilitassem concluir a veracidade (ou não) de uma conclusão a partir de premissas contidas em situações-problema descritas no instrumento. O teste também possibilita que o participante possa responder se não existe informação suficiente nas situações-problema que permita concluir algo, ou seja, esta habilidade requer do participante reflexão profunda sobre o argumento existente nas situações-problema do teste. Abaixo observa-se um exemplo de item existente no teste:

Exemplo: Seu João organizou um passeio e todos os seus sobrinhos foram ao passeio.

Você pode concluir que (assinale V, P ou F):

(V) (P) (F) Todos que foram ao passeio são sobrinhos do seu João.

Legenda: V necessariamente verdadeiro
P não é possível ter certeza
F necessariamente falso

Este instrumento adota o modelo cognitivo chamado *Transitive Chain Theory*, de Guyote e Sternberg (1981), que se constitui numa teoria acerca do raciocínio silogístico. Esta teoria se propõe a entender os processos de raciocínio silogístico (na qual uma conclusão pode ser retirada de premissas) a partir de um percurso de pensamento e de linguagem, ao mesmo tempo que também se constitui num movimento de raciocínio, permitindo a dedução de algo. As informações contidas nas premissas formariam assim um conjunto de relações, as quais seriam integradas no raciocínio graças a um conjunto de regras de cadeias transitórias existentes nas relações estabelecidas pelo argumento silogístico. Esta teoria defende a utilização de tais regras possibilita a criação de novos componentes de informação capazes de se combinar formando uma representação integrada de toda a informação existente no argumento.

Desta forma entende-se que a representação simbólica dos componentes de informação torna possível a representação da relação entre dois conjuntos que estão sendo analisados nas premissas do silogismo. De forma mais simples, esta teoria descreve uma construção de representações simbólicas as quais formariam nódulos de informação, os quais por sua vez formariam adequadas conexões entre si utilizando a memória. Assim, tal representação simbólica

auxiliaria e tornaria adequada a possibilidade de formação de processos dedutivos, os quais são exigidos para que o raciocínio silogístico seja formado.

Guyote e Sternberg (1981) realizaram pesquisas e concluíram que esta teoria fornece informações adequadas para silogismos que utilizam quantificadores e que possuam relações lógicas do tipo categórico e do tipo condicional. Seus estudos também defendem que a *Transitive Chain Theory* responde bem a questões inerentes ao raciocínio silogístico, tais como: a representação da informação contida nas premissas, a possibilidade de combinação de informações existentes nas premissas, entender a origem das dificuldades do raciocínio silogístico e, por fim, entender melhor a generalidade existente nos processos que são utilizados no raciocínio silogístico. Diante disto compreende-se o porquê da importância desta teoria para a criação do teste de raciocínio lógico-dedutivo.

O teste de raciocínio lógico-dedutivo (Primi & Muller, não publicado) se propõe a avaliar a capacidade de análise de informações complexas existentes em uma situação-problema, relacionando-as para a produção de uma conclusão lógica diante de situações novas. Fundamentalmente, o teste apresenta problemas silogísticos que envolvem a obtenção de conclusões a partir de uma única premissa (inferência imediata), problemas com oposição entre as premissas e problemas contendo conclusões fundamentadas em duas ou mais premissas, bem como silogismos encadeados (inferências mediatas). O instrumento é formado por vinte e cinco itens, dos quais sete possuem quatro subitens cada um, de forma que o instrumento completo requer quarenta e seis respostas do participante. Todos os itens do teste são de múltipla escolha e a forma de pontuação é dicotômica (certo ou errado). Desta forma o participante poderá apresentar como pontuação mínima o valor 0 e como pontuação máxima o valor 46.

Como já relatado anteriormente, esta pesquisa toma o conceito de pensamento crítico desenvolvido por Rivas e Saiz, o qual envolve vários aspectos do funcionamento cognitivo, sendo ele mais amplo do que o próprio raciocínio lógico-dedutivo e desta forma existiu a necessidade de optar pelo uso de um teste como um indicador da avaliação do pensamento crítico, sendo então escolhido este teste que possui especificidade na avaliação do raciocínio lógico-dedutivo. Percebe-se também que existem poucos instrumentos validados para a avaliação do pensamento crítico na literatura internacional e apesar da existência de um teste de raciocínio lógico-dedutivo que foi criado para a realidade brasileira, este ainda não foi estudado com enfoque nas suas propriedades psicométricas, ou seja, é possível que sejam encontradas algumas informações estatísticas inerentes

ao contexto cultural brasileiro, com possibilidade de implicações que o pensamento crítico parece ter nas tomadas de decisão e práticas educacionais.

Assim, o objetivo deste trabalho foi o de investigar as propriedades psicométricas do Teste de Raciocínio Lógico-Dedutivo (Primi & Muller, não publicado) por meio de duas grandes teorias psicométricas: a primeira seria a partir da Teoria Clássica dos Testes (TCT) buscando evidências de validade baseada na estrutura interna (validade fatorial) e evidências de validade com base nas relações com variáveis externas (raciocínio abstrato e os cinco grandes fatores de personalidade), e, por último investigar os índices de fidedignidade dos fatores que vierem a ser encontrados na análise fatorial. A segunda teoria utilizada foi a Teoria de Resposta ao Item (TRI), buscando-se obter análise do ajuste dos dados ao Modelo de Rasch da TRI (*infit* e *outfit*); Analisar a adequação do nível de dificuldade dos itens (parâmetro b) ao nível de habilidade dos sujeitos θ (Θ) e por fim analisar o espectro de precisão do instrumento. Foi escolhida uma amostra de população da região metropolitana do Recife para aplicação dos testes de Raciocínio Lógico-Dedutivo (Primi & Muller, não publicado); aplicação da Prova de Raciocínio Abstrato (RA), Forma B, da Bateria de Provas de Raciocínio (BPR-5) para fins de coleta de dados acerca da inteligência fluida (capacidade de raciocínio diante de situações novas, originando novas definições e entendendo as consequências do raciocínio) e por fim a aplicação da Bateria Fatorial de Personalidade (BFP), a qual avalia a personalidade utilizando o modelo dos cinco grandes fatores (neuroticismo, extroversão, responsabilidade, socialização e abertura à experiência).

Esses estudos serão apresentados a seguir no formato de dois artigos. O primeiro, intitulado *Estudos de validade de um teste de raciocínio lógico-dedutivo*, versa sobre as evidências de validade baseadas na estrutura interna do teste de Raciocínio lógico dedutivo, bem como sobre a avaliação da precisão (fidedignidade) dos fatores encontrados. Por sua vez o segundo artigo, intitulado *Análise de um teste de raciocínio lógico-dedutivo com TRI*, versa sobre a investigação das propriedades psicométricas do teste de raciocínio lógico dedutivo por meio da Teoria de Resposta ao Item (TRI) com objetivo de analisar o ajuste dos dados ao Modelo de Rasch da TRI (*infit* e *outfit*), como também verificar a adequação do nível de dificuldade dos itens (parâmetro b) ao nível de habilidade dos sujeitos θ (Θ) e por fim discutir o espectro de precisão do instrumento.

REFERÊNCIAS

- Almeida, L.S; Guisande, M.A; Pereira, A; Joly, M.C.R.A; Donaciano, B; Mendes, T; Ribeiro, M.S. (2009). Escala de competências de estudo (ECE-Sup): fundamentos e construção. In: B. D. Silva, L. S. Almeida, A. Barca. X. M. Peralbo (Orgs), *Atas do X Congresso Internacional Galego-Português de Psicopedagogia* (pp.4282-4292). Braga: CIED – Universidade do Minho.
- Andrade, D.F.; Tavares, H.R.; Valle, R.C. (2000). *Teoria de Resposta ao Item: Conceitos e aplicações*. São Paulo, SP: Associação Brasileira de Estatística.
- Andrade D.F., Bortolotti S.L.V. (2007). Aplicação de um Modelo de Desdobramento Graduado Generalizado- GGUM da Teoria da Resposta ao Item. *Estudos Avaliação Educ.* 18(37):157-87.
- Andriola, W.B. (1998) Utilização da Teoria da Resposta ao Item (TRI) para a organização de um banco de itens destinados a avaliação do raciocínio verbal. *Psicologia Reflexão e Crítica*. Porto Alegre. Vol.11, n.2, p.295-308.
- Araújo, E.A.C; Andrade, D.F; Bortolotti, S.L.V. (2009). Teoria de resposta ao item. *Rev. Esc. enferm USP*. 43(Esp):1000-8.
- Balbinotti, Marcos Alencar Abaide. (2005). Para se avaliar o que se espera: reflexões acerca da validade dos testes psicológicos. *Aletheia*, Canoas-RS, N.21, jan-jun, pp.43-52.
- Barnes, C.A. (2005). Critical thinking revisited: Its past, present, and future. *New Directions for Community Colleges*, Summer 2005, pp. 5-13.
- Baumgart, Viviane de Oliveira; Nascimento, Elizabeth do. (2004). A bateria de provas de raciocínio lógico (BPR-5) aplicada a um contexto organizacional. *Psico-USF*, v.9, n.1, pp.1-10.
- Bernard, R.M.; Zhang, D.; Abrami, P. C.; Sicol, F.; Borokhovski, E.; Surkes, M. A. (2008). Exploring the structure of the Watson–Glaser Critical Thinking Appraisal: One scale or many subscales? *Thinking Skills and Creativity*, v.3(1), pp.15-22.
- Bernardes, M. E. M; (2011). O pensamento na atividade prática: implicações no processo pedagógico. *Psicologia em estudo*, Maringá, v. 16, n.4. pp.521-530, out/dez.
- Bock, Ana Mercês Bahia; Furtado, Odair & Teixeira, M. L. T. (1999). *Psicologias: uma introdução ao estudo da psicologia*. São Paulo: Saraiva, 18 – 29; pp.150-164; pp.214-228.
- Bloom, B.S. (1956). Taxonomy of educational objectives: Handbook I. *The cognitive domain*. New York: McKay.

- Brady, M. (2008). Cover the material: or teach students to think? *Educational Leadership*, 65, pp.64-67.
- Brookfield, S.D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting*. San Francisco: Jossey-Bass.
- Carelli, M.J.G; Santos, A.A.A. (1998). Condições temporais e pessoais de estudo em universitários. *Psicologia escolar e educacional*, 2, pp.265-278.
- Carroll, R. (2005). *Becoming a critical thinker: A guide for the new millennium (2nd ed.)*. Boston: Pearson Custom Publishing.
- Castelar, Marilda. (1992). *Os multimeios e os instrumentos de seleção de pessoas*. Dissertação de mestrado apresentada ao departamento de multimeios do instituto de artes da Unicamp. Campinas: pp. 10-34.
- De Chiaro, S. & Leitão, S. (2005). O papel do professor na construção discursiva da argumentação em sala de aula. *Psicologia: Reflexão e Crítica* (UFRGS. Impresso), Porto Alegre, v. 18, n. 3, pp. 350-357.
- Embretson, S.E.; Reise, S.P. (2000). *Item Response Theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J.B. Baron & R.J. Sternberg (Eds), *Teaching thinking skills*. Freeman and company. New York, pp. 9-26.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and need research. *Educational Researcher*, 18, pp.13-16.
- Ennis, R.H. (2003). Critical thinking assessment. In D. Fasko (Ed.), *Critical thinking and reasoning. Current Research, theory, and practice*. (pp.293-313). Cresskill, NJ: Hampton Press.
- Fernández, J.M.. (1990). *Teoría de Respuesta a los Ítems. Un nuevo enfoque en la evolución psicológica y educativa*. Madrid: Ediciones Pirámide S.A.
- Filho, Marcilio Lira de Souza; Belo, Raquel; Gouveia, Valdiney Veloso. (2006). Testes psicológicos: Análise da produção científica brasileira no período 2000-2004. *Psicologia, ciência e profissão*, 26(3), pp.478-489.
- Fletcher, P. (1994). A teoria da resposta ao item: Medidas invariantes do desempenho escolar. *Ensaio: Avaliação e Políticas Públicas em Educação*, 1(2), pp.21-28.
- Gardner, Howard. (1995) *Inteligências múltiplas – a teoria na prática*. Porto Alegre: Artmed, pp. 12-18.

- Gardner, H.; Kornhaber, M.; Wake, W. K. (1998). *Inteligência, múltiplas perspectivas*. Porto Alegre: ArtMed, pp. 213-225.
- Guyote, M.J; Sternberg, R.J. (1981). *A transitive-chain Theory of Syllogistic Reasoning*. *Cognitive psychology*, 13, pp. 461-525.
- Greeno, J. G. (1989). A Perspective on Thinking. *American Psychologist*, Vol. 44, n.2, pp.134-141.
- Halonen, J. S. (1995). Desmystifying critical thinking. *Teaching of Psychology*, 22:1, pp.75-81.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American psychologist*, 53(4), pp.449-455.
- Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking (4th ed.)*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Halpern, D. F. (2006). *Halpern Critical Thinking Assessment Using Everyday Situations: Background and scoring standards*. Unpublished report.
- Halpern, D. F. (2010). *Manual HCTA, Halpern Critical Thinking Assessment (Version 2.1)*. Mödling, Áustria: Schuhfried.
- Junker W.; Sijtsma K.; (2006). Item Response Theory: past performance present, developments, and future expectations. *Behaviormetrika*. 33(1):pp.75-102.
- Kopnin, P.V. (1978). *A dialética como lógica e teoria do conhecimento*. Tradução: Paulo Bezerra. Rio de Janeiro: Civilização Brasileira.
- Leitão, S. (1996). Perspectivas no estudo da argumentação quotidiana. *Psicologia. Teoria e Pesquisa*, Brasília, v. 12, n. 1, pp.11-21.
- Leitão, S. (2007). Argumentação e desenvolvimento do pensamento reflexivo. *Psicologia: Reflexão e Crítica* (UFRGS. Impresso), v. 20, pp. 454-462.
- Lindner, R. W; Harris, B. (1993). Self-regulated learning: Its assessment and instructional implications. *Educational Research Quarterly*, 16, pp.29-37.
- McPeck, J.E. (1981). *Critical thinking and educations*. New York: St. Martins. pp. 05-20.
- Mentkowsiki, M. (1991). *Higher education assessment and national goals for education: Some issues, assumptions, principles for discussion* (American Psychological Association's Task Force Report on Psychology in Education and the Mid-Continent Regional Educational Laboratory). Racine, W: Wingspread.

- Meece, J. L. (1994). The role of motivation on self-regulated learning. In: D.H. Shunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, NJ: Erlbaum.
- Minnich, E. (1990). *Transforming knowledge*. Philadelphia: Temple University Press. pp.40-65.
- Muñiz, J. & Hambleton, R.K. (1992). Melo siglo de Teoría de Respuestas a los Ítems. *Anuario de Psicología*, 52, 41-66.
- Muñiz, J. (2004). La validación de los tests. *Metodología de las ciencias del comportamiento*, 5, 121-141.
- Nieto, A.M.; Saiz, C.; Orgaz, B. (2009). Análisis de la propiedades psicométricas de la versión española del HCTAES-Test de Halpern para la evaluación del pensamiento crítico mediante situaciones cotidianas. *Revista electrónica de metodología aplicada*, v.14, n.1, pp.1-15.
- Nolt, J & Rohatyn, D. (1991). *Lógica*. Tradução Mineko Yamashita. São Paulo: McGraw-Hill. pp. 01-65.
- Noronha, Ana Paula Porto. (2003) Estudos de validade e de precisão em testes de inteligência. *Paidéia*, 13(26), pp.163-169.
- Noronha, A.P.P.; Vendramini, C.M.M; Canguçu, C.; Souza, C.V.R; Cobêro, C.; Paula, L. M.; Franco, M.O; Lima, O.M.P.; Guerra,P.B.C.; Filizatti, R. (2003). Propriedades psicométricas apresentadas em manuais de testes de inteligência. *Psicologia em estudo*, Maringá, V. 8, n. 1, pp. 93-99.
- Noronha, Ana Paula Porto; Alchieri, João Carlos. (2004). Conhecimento em avaliação psicológica. *Rev. Estudos de psicologia, PUC-Campinas*, v. 21, n. 1, pp. 43-52.
- Noronha, Ana Paula Porto; Primi, Ricardo; Alchieri, João Carlos. (2005). Instrumentos de avaliação mais conhecidos/utilizados por psicólogos e estudantes de psicologia. *Psicologia: Reflexão e crítica*, 18(3), pp. 390-401.
- Nunnally, Jum C. (1970). *Introduccion a la medicion psicologica*. Buenos Aires: Editorial Paidós.
- Paul, R. (2005). The state of critical thinking today. *New directions for community colleges*, summer 2005, pp.27-38.
- Pasquali, L. (2003). *Psicometria: Teoria dos Testes na Psicologia e na Educação*. Petrópolis: Vozes. pp. 67-105.
- Pasquali, L.; Primi, R. (2003). Fundamentos da Teoria de Resposta ao Item – TRI. *Avaliação psicológica*, 2(2), 99-110.

- Pintrich, P.R; Smith, D.A; Garcia, T; Mckeachie, W.J. (1993). Reability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement*, 53, pp.801-813.
- Primi, R. (2002). Inteligência Fluida: Definição Fatorial, Cognitiva e Neuropsicológica. *Paidéia*, 12, pp.57-75.
- Primi, R. (2010). Avaliação psicológica no Brasil: Fundamentos, situação atual e direções para o futuro. *Psicologia: Teoria e Pesquisa*, vol.26, n. especial, pp.25-35.
- Rivas, Silvia F. & Saiz, Carlos (2012). Validación y propiedades psicométricas de La prueba de pensamiento crítico – PENCRISAL. *Revista Electrónica de Metodologia Aplicada*. Vol.17, n 1, pp.18-34.
- Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, 65, pp.1-10.
- Saiz, Carlos. (2002) *Pensamiento crítico – Conceptos básicos y actividades prácticas*. Psicología pirâmide, pp.16-28.
- Saiz, C. & Rivas, S. F. (2010). Melhorar El pensamiento critico contribuye al desarrollo personal de lós jóvenes? In: H. J. Ribeiro & J. N. Vicente (Eds.), *O lugar da lógica e da argumentação no ensino da filosofia*. Coimbra: Unidade I & D, Linguagem, Interpretação e Filosofia. pp. 39-52.
- 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), pp.241-250.
- Spearman C. (1904). “General Intelligence”, objectively determined and measured. *Am J Psychol*. 15(2), pp.201-93.
- Thurstone L.L.; (1927). A Law of comparative judgment. *Psychol Rev*. 34(2), pp.273-86.
- Thurstone L.L. (1928). Attitudes can be measured. *Am J Sociol*. 26(2), pp.249-69.
- Urbina, S. (2007). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artmed. pp. 121-211.
- Vasconcelos, R.M; Almeida, L.S; Monteiro, S.C. (2005). Métodos de estudo em alunos do 1º ano da universidade. *Psicologia escolar e educacional*, 9, pp.195-202.
- Vendramini, C.M.M., Silva, M.C; Canale, M. (2004). Análise de itens de uma prova de raciocínio estatístico. *Psicologia em Estudo*, 9(3), pp.487-498.
- West, R.F.; Toplak, M.E & Stanovich, K.E., (2008). Heuristics and biases as measures of critical thinking: associations with cognitive abilities and thinking dispositions. *Journal of educational psychology*, 100(4), pp.930-941.

- Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist*, 30, pp.173-187.
- Yanchar, Stephen C.; Slife, Brent D. & Warne, Russell. (2008). Critical Thinking as Disciplinary Practice. *Review of General Psychology*, Vol. 12, n.3, pp. 265–281.
- Zimmerman, B.J. (2000). Self-efficacy: An essential motive to learn. *Contemporary educational psychology*, 25, pp.82-91.

Amorim, Mozart. (2015). Estudos de validade de um teste de raciocínio lógico-dedutivo. Dissertação de Mestrado, Pós-Graduação em Psicologia Cognitiva, Universidade Federal de Pernambuco, Recife.

RESUMO

Esta pesquisa pretendeu investigar as propriedades psicométricas de um Teste de Raciocínio Lógico-Dedutivo (TRLD). A amostra foi composta por 192 participantes predominantemente do sexo feminino (65,6%), todos com idades superiores a 18 anos (média igual a 23,9 e desvio padrão de 9,33). Para investigação da validade fatorial foi empregada uma análise fatorial exploratória, pela qual foram encontrados cinco fatores primários e dois de segunda ordem, que, no geral, apresentaram bons índices de consistência interna. Esses fatores se correlacionaram significativamente com uma medida de raciocínio abstrato (validade convergente) e pobremente com traços de personalidade (validade discriminante). O instrumento apresenta boas propriedades psicométricas para ser recomendado para uso em pesquisas.

Palavras-chave: Pensamento crítico; avaliação psicológica; testes psicológicos; lógica informal.

Amorim, M. (2015). Validity Studies of a logical-deductive reasoning test. Master's degree thesis, Cognitive Psychology Graduate School, Universidade Federal de Pernambuco, Recife.

ABSTRACT

This research aimed to investigate the psychometric properties of a Reasoning Test Logical-Deductive (RTLTD). The sample consisted of 192 participants predominantly female (65.6%), all aged over 18 years (mean of 23.9 and standard deviation of 9.33). To investigate the factorial validity was used exploratory factor analysis, for which there are five primary factors and two second order, which, in general, showed good internal consistency indices. These factors significantly correlated with a measure of abstract reasoning (convergent validity) and poorly with personality traits (discriminant validity). The instrument has good psychometric properties to be recommended for use in research.

Keywords: Critical thinking; psychological assessment; psychological tests; informal logic

INTRODUÇÃO

Os estudos em lógica compreendem a investigação dos métodos e princípios que regem a distinção do raciocínio correto do incorreto, facilitando desta maneira a compreensão dos argumentos. A lógica também é comumente conceituada como “a ciência das leis do pensamento” (Copi, 1974, p. 20), sendo esta uma terminologia pouco aceita por seus estudiosos, uma vez que outras ciências, como a psicologia, também estudam as leis que regem o pensamento. Também não é considerada muito completa a concepção de que a lógica é a “ciência do raciocínio”, uma vez que os especialistas em lógica não estão interessados nos diversos percursos que a mente perpassa para concretizar determinado raciocínio, assim a pergunta central para um estudioso da lógica poderia ser: a conclusão a que se chegou deriva das premissas (conjunto de uma ou mais sentenças declarativas) usadas ou pressupostas? Desta maneira, a lógica classificará o argumento (conjunto de enunciados que se relacionam) como correto ou incorreto, sendo este o seu problema central, com o consequente desenvolvimento de métodos e técnicas para fins de tal detecção (Copi, 1974; Nolt & Rohatyn, 1991).

Tradicionalmente, a ideia da lógica remonta à capacidade de validar ou não os argumentos, sem que se diga algo a respeito do assunto, permitindo assim neutralidade quanto ao tema em questão. Assim, os sistemas em que se pode aplicar regras de raciocínio (independentemente do seu tema), serão considerados objeto de estudo lógico. Dentre os elementos formadores de uma estruturação lógica, têm-se as premissas, que são sentenças declarativas que sustentam (argumento válido) ou não (argumento inválido) uma conclusão ou sentença conclusiva. Assim depreende-se que o termo “argumento” em lógica se constitui na articulação entre as premissas gerando uma conclusão (Haack, 2002; Nolt & Rohatyn, 1991).

Para a filosofia, o elemento lógico (capacidade de entender a existência ou não de relacionamento entre os enunciados) não deve assumir uma função de tradução de uma intuição mental, mas sim uma função de validação a partir de uma demonstração sistematizada, a qual permitirá a intelecção do problema abordado com sua posterior resolução (Gueroult, 2007). Ainda de acordo com a classificação empregada pela filosofia para a temática da lógica, pode-se encontrar duas amplas classificações: a lógica formal e a lógica informal.

Segundo Pinto (1969), a lógica formal se presta a identificar as características externas dos objetos a partir da constituição do pensamento, buscando relações superiores do pensar a partir de um conteúdo puramente formal, permitindo assim que o processo dialético forneça reflexão suficiente a ponto do ser pensante conhecer e também explicar o ambiente que o cerca, entendendo-se também como um elemento constituinte e transformador deste meio. A lógica formal se fundamenta desta maneira num tipo de linguagem não-comum ou de linguagem simbólica, preocupando-se basicamente com a estrutura do raciocínio (relação entre os conceitos) e desta forma procurando entender a relação que possa existir entre as premissas e a conclusão. Contudo, nem sempre a lógica formal consegue dar conta de tal objetivo, devendo assim o pesquisador utilizar outros tipos de abordagem da lógica para tentar responder suas indagações acerca do objeto em estudo (Bernardes, 2011). Dentre os componentes inerentes ao estudo da lógica temos principalmente dois tipos de argumentação: a dedução (inferência que parte do geral para o particular) e a indução (inferência que parte do particular para o geral) (Mundim, 2002).

A lógica informal se estrutura a partir de formas cotidianas de raciocínio, fugindo de certo modo da modalidade da lógica formal, assumindo assim caráter mais indutivo do que dedutivo (Ennis, 1996). Na década de 60 pesquisadores ampliaram os estudos desta área iniciando desta maneira as pesquisas sobre o chamado pensamento crítico e desta maneira, o pensamento crítico abarca não apenas as formas de raciocínio formais, mas também as informais, permitindo investigações em tipos de manifestações reflexivas deixadas de lado pela lógica formal, mas que são igualmente importantes na construção do raciocínio, sem deixar de lado capacidades fundamentais do raciocínio propostas, como habilidades básicas de comprovação de hipóteses, resolução de problemas e tomada de decisão, produzindo assim testes e avaliações no intuito de melhor entender o pensamento crítico (Saiz, 2002). Assim o pensamento crítico pode ser enquadrado na área de estudos da cognição humana, funcionando através de processos mentais amplos, profundos e complexos, exigindo interpretação, reflexão e tomada de decisão do indivíduo.

Dentre os instrumentais existentes para a avaliação do pensamento crítico pode-se citar o *Halpern Critical Thinking Assessment Using Everyday Situations* – HCTAES, o qual pretende avaliar cinco habilidades relacionadas ao pensamento crítico: a análise do argumento, a comprovação de hipóteses, o raciocínio verbal, a probabilidade e incerteza e tomada de decisão (Halpern, 2006). Nieto, Saiz e Orgaz (2009) encontraram uma medida de adequação da amostra através da prova de Kaiser-Mayer-Olkin com valor de 0,70 e prova de esfericidade de Bartlett

($\chi^2=2214,44$, $p<0,05$), A partir disso, chegaram a conclusão de que neste teste não existia uma única dimensão avaliada, sugerindo que o pensamento crítico seja formado por um conjunto de habilidades representadas pelos fatores encontrados. Os índices de fidedignidade desse fatores, no entanto, medidas pelo Alfa de Cronbach foram de 0,635 para a habilidade de comprovação de hipóteses, 0,367 para a habilidade de raciocínio verbal, 0,465 para a análise do argumento, 0,340 para probabilidade e incerteza e por fim o valor de 0,487 para a habilidade de tomada de decisão, contudo estes valores obtidos não foram considerados adequados aos valores mínimos exigidos por critérios internacionais (Urbina, 2007).

Outro teste que avalia o pensamento crítico é o *Watson-Glaser Critical Thinking Assessment* (WGCTA), o qual possui uma bateria de 80 itens com resposta múltipla, avaliando as seguintes dimensões: interpretação, reconhecimento de pressupostos, avaliação de argumentos, dedução e inferências (West, Toplak & Stanovich, 2008). Bernard et al (2008) realizaram pesquisa correlacionando as cinco dimensões avaliativas do WGCTA e tiveram como resultado correlações estatisticamente significativas, moderadas e com pouca heterogeneidade, o que os permitiu concluir que o teste é capaz de avaliar uma competência geral, mas que em se tratando das dimensões já citadas o teste não conseguia captá-las a ponto de gerar interpretações seguras acerca do desempenho do participante em cada dimensão avaliada.

Rivas e Saiz (2012) desenvolveram outro instrumento para avaliação do pensamento crítico chamado PENCRIAL. Esse instrumento é baseado em três aptidões principais: o raciocínio, a resolução de problemas e a tomada de decisão, sempre mantendo em mente que o instrumental fosse capaz de avaliar a utilização do pensamento para solucionar questões corriqueiras. O PENCRIAL pretende avaliar distintos domínios cognitivos: raciocínio indutivo, raciocínio dedutivo, raciocínio prático, tomada de decisão e a resolução de problemas. Numa pesquisa para validação da versão espanhola do PENCRIAL os autores encontraram para a consistência interna dos 35 itens (Alfa de Cronbach) o valor de 0,632, considerado dentro de níveis estatísticos aceitáveis. A precisão entre avaliadores obteve valores de Kappa dentro do intervalo de 0,6 a 0,9, com $p<0,001$.

Por sua vez o teste de raciocínio lógico-dedutivo (Primi & Muller, não publicado) foi desenvolvido objetivando avaliar a inteligência fluída baseada na capacidade de análise de informações existentes num contexto novo de uma situação-problema, combinando-as para a produção de uma conclusão lógica. Este teste consta com quarenta e seis itens e subitens com

resposta de múltipla escolha e forma de pontuação dicotômica (certo ou errado). Assim o participante apresentará como pontuação mínima o valor 0 e como pontuação máxima o valor 46.

Embora nesta pesquisa este teste seja o único que possa servir como indicador do nível de habilidade do pensamento crítico, na literatura pesquisada não foi encontrada até a presente data* nenhuma pesquisa que investigasse suas propriedades psicométricas, apenas alguns trabalhos utilizaram este teste, porém suas propriedades psicométricas não se constituíram em objeto de análise. Diante disso, o objetivo desta pesquisa foi investigar as propriedades psicométricas do Teste de Raciocínio Lógico-Dedutivo (TRLDD), com base na Teoria Clássica dos Testes. Mais especificamente, os objetivos são: 1) buscar evidências de validade baseadas na estrutura interna do instrumento; 2) Verificar os coeficientes de fidedignidade dos fatores que vierem a ser encontrados; 3) buscar evidências de validade baseadas na relação com variáveis externas, neste caso, uma medida tradicional de inteligência fluída e medidas dos cinco grandes fatores de personalidade.

* Foram pesquisados os sites de periódicos da SciELO (www.scielo.org), Periódicos CAPES (www.periodicos.capes.gov.br), BVS-PSI (www.bvs-psi.org.br), biblioteca digital da UNICAMP (www.bibliotecadigital.unicamp.br), biblioteca digital brasileira de teses e dissertações (<http://bdtd.ibict.br/>), biblioteca de teses e dissertações da UNICAP (www.unicap.br/tede/tde_busca), e por fim o GOOGLE Acadêmico. O período de pesquisa foi compreendido entre 01.10.2014 a 15.01.2015.

MÉTODO

Participantes

Foram convidados através de email, bem como em salas de aula do curso de psicologia da UFPE. A participação em todas as etapas de coleta de dados ocorreu de forma presencial e coletiva. No total participaram da pesquisa 249 voluntários, dos quais 57 foram excluídos por apresentarem idade inferior a 18 anos, uma vez que o teste de Raciocínio abstrato, forma B foi criado para aplicação em adultos e desta forma considerou-se conveniente que toda a amostra consistisse somente em participantes adultos. A amostra final utilizada para a pesquisa foi composta pelo total de 192 participantes, sendo 66 do sexo masculino (34,4%) e 126 do sexo feminino (65,6%). A idade dos participantes ficou entre 18 e 64 anos, com média de 23,9 e desvio padrão de 9,33. Em relação a etnia, 44,8% se declaram brancos, 38,5% pardos, 12,5% negros e oito participantes não responderam sobre sua etnia, correspondendo a 4,2%. Sobre a escolaridade, 1,6% dos participantes possuem o ensino fundamental, 38% possuem o ensino médio, 43,2% possuem nível superior de ensino, 13% são pós-graduados e 4,2% não informaram a escolaridade. No quesito renda, 45,3% afirmaram que possuem renda até R\$ 1.700,00, enquanto 21,9% afirmaram ter renda entre R\$ 1.700,00 e R\$ 3.400,00. Os que assinalaram possuir renda entre R\$ 3.400,00 e R\$ 5.100,00 constituíram 10,4% da amostra e 17,7% afirmaram pertencer ao grupo com renda acima de R\$ 5.100,00. Nove participantes não informaram sua renda, constituindo 4,7% da amostra.

Todos os participantes responderam ao teste de raciocínio lógico-dedutivo, mas apenas 95 pessoas responderam adicionalmente a Forma B da Prova de Raciocínio Abstrato (RA) da Bateria de Provas de Raciocínio (BPR-5) e a Bateria Fatorial de Personalidade (BFP).

Instrumentos

Foi utilizado um questionário de informações pessoais para descrição da amostra, preenchidos por cada participante antes da aplicação dos testes, os quais continham perguntas acerca de dados sócio-econômicos dos participantes. Além disso, foram utilizados os seguintes instrumentos: Teste de Raciocínio Lógico-Dedutivo (Primi & Muller, não publicado), Prova de Raciocínio Abstrato (RA), forma B, da Bateria de Provas de Raciocínio (BPR-5) (Primi & Almeida, 2000) e a Bateria Fatorial de Personalidade (BFP) (Nunes, Hutz & Nunes, 2010), que são descritas a seguir.

O Teste de Raciocínio Lógico-Dedutivo (Primi & Muller, não publicado), avalia a capacidade de analisar informações de uma situação-problema, combinando-as para a produção de conclusões lógicas. O instrumento é constituído por 25 itens, dos quais 7 itens possuem 4 subitens cada um, formando assim 28 novas perguntas. Assim o instrumento apresenta 46 respostas solicitadas ao participante no teste como um todo. A forma de escolha da resposta varia dependendo do item, sendo de múltipla escolha (até 5 possíveis) em 9 itens e possuindo apenas 3 opções de escolha em outros 9 itens distintos. Por sua vez o teste ainda tem 28 subitens possuindo apenas três possibilidades de escolha da resposta correta. Já a forma de pontuação para cada item é do tipo dicotômica (certo ou errado). Desta forma o participante poderá apresentar como pontuação mínima o valor 0 e como pontuação máxima o valor 46.

A prova de Raciocínio Abstrato (RA), Forma B, da Bateria de Provas de Raciocínio (BPR-5), é composta por 25 itens, envolvendo analogias com figuras geométricas. Os escores foram as somas dos acertos nessa prova. Este instrumento apresentou bons indicadores de validade e precisão para a população brasileira e foi aplicado e avaliado conforme recomendações constantes no manual (Primi & Almeida, 2000).

Bateria Fatorial de Personalidade (BFP), que é baseada no Modelo dos Cinco Grandes Fatores de Personalidade (neuroticismo, extroversão, responsabilidade, socialização e abertura à experiência). Possui 126 itens, para serem respondidos por meio de uma escala do tipo Likert de sete pontos, na qual 1 representa "Absolutamente não me identifico com a frase" e 7 significa "Descreve-me perfeitamente". A prova, que apresentou bons indicadores de validade e precisão

para a população brasileira foi aplicada e avaliada conforme recomendações constantes no manual (Nunes, Hutz & Nunes, 2010).

Procedimentos

O projeto foi submetido ao comitê de ética para pesquisas com seres humanos da Universidade Federal de Pernambuco e somente após sua aprovação foi iniciada a fase de coleta de dados. Os participantes foram informados quanto aos objetivos e procedimentos envolvidos na pesquisa. Essas informações também constavam em um Termo de Consentimento Livre e Esclarecido (TCLE), que todos os voluntários da pesquisa leram e concordaram em assinar, permitindo sua participação nesta pesquisa bem como a utilização dos dados coletados de suas respostas para fins estatísticos, resguardando desta forma suas informações particulares.

Os dados foram coletados a partir de apresentação dos testes impressos com preenchimento de respostas em uma folha específica, chamada folha de respostas personalizada para cada tipo de teste. Os participantes foram contatados em salas de aula na Universidade Federal de Pernambuco bem como através de email, recebendo um convite para participação nesta pesquisa.

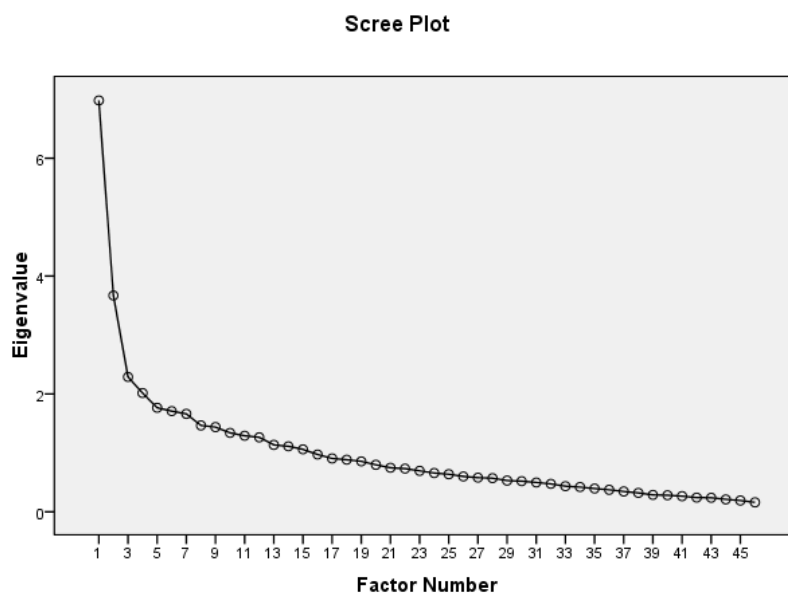
As respostas foram digitadas numa planilha eletrônica e a análise dos dados pretendeu verificar os seguintes pontos: 1) Para a busca de evidências de validade baseadas na estrutura interna foi empregada uma análise fatorial, esperando-se que se formasse um fator geral, relacionado com o raciocínio lógico-dedutivo; 2) Para investigar a fidedignidade do instrumento foram calculados os coeficientes de Kuder-Richardson dos fatores obtidos na análise fatorial. Esperava-se que esses coeficientes fossem superiores a 0,7; 3) Para investigar a relação com variáveis externas foram empregados os coeficientes de correlação de Pearson. Esperava-se que os fatores de raciocínio lógico-dedutivo apresentassem correlações moderadas, positivas e significativas com o raciocínio abstrato, de modo a apoiar a validade convergente com medidas tradicionais de inteligência e, ao mesmo tempo, mostrar que existe alguma variância única desses fatores, caracterizando a sua diferenciação da medida de raciocínio abstrato. Além disso, esperava-se que as correlações dos fatores de raciocínio lógico-dedutivo fossem baixas e não significativas

com traços de personalidade, apoiando a sua discriminância desse construto. Os dados foram analisados com auxílio dos softwares MPLUS Versão 6.11 (Muthen & Muthen, 2010) e SPSS versão 20.

RESULTADOS

Os dois primeiros objetivos deste trabalho foram buscar evidências de validade baseadas na estrutura interna do instrumento e avaliar a precisão (fidedignidade) dos fatores encontrados. Para isso foi realizada uma análise fatorial e calculados os coeficientes de fidedignidade de Kuder-richardson, respectivamente. Em primeiro lugar foi investigado o número de fatores a serem retidos na análise, por meio da inspeção do gráfico de sedimentação. Embora fosse esperado apenas um fator relacionado ao processamento lógico-dedutivo de uma forma geral, a figura 1 apresenta o gráfico de sedimentação (*scree-plot*), mostrando que surgiram cinco fatores considerados como mais importantes.

Figura 1: Gráfico de sedimentação (*scree-plot*)



Com base nisso, uma análise fatorial exploratória para dados categóricos foi implementada, com auxílio do *software* MPlus 6.11, para extração dos cinco fatores, por máxima verossimilhança e rotação *geomin*. A Tabela 1 mostra as cargas fatoriais dessa análise.

Tabela 1 - Cargas fatoriais dos cinco principais fatores

Itens	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5
U1	0.377	0.399	0.205	0.041	0.141
U2	0.136	-0.200	-0.146	0.227	0.643
U3	0.354	0.369	0.223	0.271	-0.037
U4	0.733	-0.008	0.144	-0.117	0.104
U5	0.011	0.668	-0.251	0.031	-0.141
U6	0.341	0.196	0.046	0.055	0.376
U7	-0.114	0.456	0.323	0.214	-0.115
U8	0.496	-0.080	0.447	0.082	-0.298
U9	-0.063	0.646	-0.095	-0.072	-0.204
U10	-0.218	0.471	-0.335	-0.007	0.083
U11	-0.241	0.623	0.063	-0.007	-0.141
U12	-0.449	0.275	-0.312	0.089	0.257
U13	1.137	-0.005	-0.417	0.379	0.002
U14	0.834	0.204	0.001	0.067	0.011
U15	-0.112	-0.240	0.032	0.240	0.490
U16	0.204	0.322	-0.036	-0.065	-0.085
U17	0.200	-0.032	0.028	0.178	0.409
U18	0.441	0.274	-0.040	0.116	0.241
U19	0.092	0.376	0.009	0.746	0.063
U20	-0.041	0.337	0.121	0.755	-0.039
U21	-0.059	-0.093	0.110	0.362	0.261
U22	0.080	0.503	-0.120	0.653	-0.045
U23	0.020	0.627	0.479	-0.114	0.066
U24	0.371	-0.041	0.224	-0.211	0.365
U25	-0.055	0.571	0.337	0.033	-0.028
U26	0.666	0.013	0.103	-0.230	0.199
U27	0.023	0.600	0.363	-0.127	0.074
U28	-0.532	0.287	0.023	0.323	0.040
U29	-0.227	0.364	0.516	0.057	-0.067
U30	-0.512	0.350	0.196	0.325	0.139
U31	-0.131	-0.004	0.864	-0.069	0.069
U32	0.188	0.095	0.203	0.147	0.156
U33	-0.050	-0.010	0.567	0.216	0.320
U34	0.237	0.114	0.534	0.069	0.079
U35	0.287	0.122	0.526	-0.053	0.186
U36	0.056	0.030	0.442	0.269	0.256
U37	0.198	-0.021	0.587	0.042	0.054
U38	0.001	0.036	0.402	0.248	-0.437
U39	0.255	-0.055	0.130	0.141	0.347
U40	0.110	-0.016	0.191	0.196	-0.499
U41	0.082	-0.081	0.454	0.360	-0.001
U42	0.172	0.066	0.676	0.047	-0.456
U43	-0.023	0.055	0.172	-0.071	0.736
U44	0.317	0.408	0.002	-0.090	0.704
U45	-0.187	-0.090	0.103	0.405	-0.024
U46	0.376	0.225	0.056	0.120	0.514
KR20	0,620	0,745	0,800	0,705	0,578

KR20 – Coeficientes de Kuder-Richardson, fórmula 20.

Por se tratar de uma análise fatorial exploratória, foi utilizado como critério considerar apenas as cargas mais elevadas (marcadas em negrito) que os itens apresentaram em cada um dos cinco fatores, desprezando cargas superiores a 0,3 que possivelmente seriam encontradas no mesmo item mas em fatores distintos. Tal trabalho mostrou-se uma tarefa complexa, pois era esperado que os itens apresentassem alta correlação entre si, pois os mesmos se pretendiam a avaliar o mesmo fator que seria o raciocínio lógico-dedutivo.

Observou-se que os itens se distribuíram organizadamente por tipo de resposta: falsa (f), verdadeira (v) ou não é possível ter certeza (p), ou pelo conteúdo das questões (raciocínio dedutivo). Foi observado que no Fator 1 predominaram os itens que exigiam que o participante detectasse o argumento inválido do raciocínio, assinalando a resposta falsa (f). No total, 10 itens apresentaram carga fatorial significativa nesse fator (U4, U8, U12, U13, U14, U18, U24, U26, U28 e U30). Também foi observado que alguns itens (U12, U28 e U30), que requeriam a detecção de que não havia informação suficiente para uma conclusão, apresentaram cargas negativas nesse fator. Isto nos faz concluir que existe sustentação estatística a favor da existência de diferenças quanto ao tipo de raciocínio empregado dentro do mesmo fator, uma vez que para responder que a afirmação é insuficiente o participante precisa utilizar um tipo raciocínio que apresentou carga contrária (negativa) ao tipo de raciocínio empregado para concluir que a resposta no teste seria falsa. O coeficiente de Kuder-Richardson desse fator foi de 0,620, o que significa que atende às normas estabelecidas pelo Sistema de Avaliação de Testes Psicológicos (SATEPSI) do Conselho Federal de Psicologia (exige KR20 igual ou maior que 0,6).

No Fator 2 ficou constatada a existência de 11 itens com cargas elevadas: U1, U3, U5, U7, U9, U10, U11, U16, U23, U25, U27. Nesses itens, o participante precisava detectar que não havia informações suficientes para um raciocínio conclusivo, devendo assinalar a alternativa “não é possível ter certeza” (p). É interessante observar que quando o participante precisa formular um raciocínio para concluir que não há informações suficientes para afirmar algo, este o faz de maneira diferente em relação ao fator 1, pois não há necessidade de raciocinar para encontrar um erro no argumento, mas apenas para detectar se há informação suficiente para se concluir algo que foi afirmado pelo item, ou seja, raciocinar para detectar que o argumento é falso parece possuir forma distinta do raciocínio empregado para se analisar se o argumento possui informação suficiente. O coeficiente de Kuder-Richardson desse fator atingiu um valor considerado adequado tanto pelo SATEPSI quando por normas internacionais, chegando a 0,745.

Em relação ao Fator 3, foi encontrado o valor mais elevado entre os coeficientes de Kuder-Richardson, o qual atingiu 0,800, ficando adequado tanto pelo SATEPSI quanto por normas internacionais. Verificou-se que este fator foi composto por 12 itens (U7, U29, U31, U32, U33, U34, U35, U36, U37, U38, U41, U42), dos quais três se referiam à detecção de não haver informação suficiente para se concluir algo (U29, U41, U42) e sete foram interpretadas com conteúdo voltado ao raciocínio dedutivo (U31, U32, U33, U34, U35, U36, U37).

O Fator 4 também ficou formado predominantemente por itens voltados para detecção de que não existem informações suficientes para se concluir algo (U19, U20 e U22). Outros itens, como o U45 e o U21, que exigem que o participante extraia uma conclusão válida ou inválida, também compuseram este fator. Em relação ao coeficiente de Kuder-Richardson o fator 4 apresentou valor de 0,705, aceito tanto pelo SATEPSI como por normas internacionais.

Por fim, observa-se que o Fator 5 foi formado pelos itens U2, U6, U15, U17, U39, U40, U43, U44, U46, possuindo semelhanças com os itens do fator 1. Entretanto, no fator 5 houve itens tanto para a detecção de raciocínios válidos (v), quanto inválidos (f). Também ocorreram itens que requeriam a detecção de que não havia informação suficiente para uma conclusão, que apresentaram cargas negativas. Nesse caso, no entanto, o coeficiente de Kuder-Richardson não atingiu valor suficiente nem pelos critérios do SATEPSI ($KR_{20}=0,6$) nem por critérios internacionais ($KR_{20}=0,7$), ficando em 0,578.

Observa-se, portanto, que os Fatores 1 e 5 são semelhantes porque neles predominam itens que exigem a detecção da adequação da resposta, seja ela verdadeira ou falsa. Já nos Fatores 2, 3 e 4, o que se observa é a predominância de itens que requeriam a detecção de que não havia informação suficiente para uma conclusão confiável. Para verificar se esses cinco fatores primários poderiam se agrupar formando novos Fatores de Segunda Ordem, foi realizada uma nova análise fatorial sobre as pontuações dos participantes nesses cinco Fatores. Essa análise resultou em dois fatores de segunda ordem, um que reuniu os fatores primários F2, F4, e F3, formando o Fator de Segunda Ordem 1 (FSO1) e outro que reuniu F1 e F5, no Fator de Segunda Ordem 2 (FSO2). Essa solução e suas respectivas cargas fatoriais encontram-se na Tabela 2.

Tabela 2 - Cargas fatoriais da Análise Fatorial de Segunda Ordem

	Fatores secundários	
	FSO1	FSO2
Fator 2	0,863	
Fator 4	0,738	
Fator 3	0,598	
Fator 1		0,777
Fator 5		0,511

No Fator de Segunda Ordem 2 há possibilidade de eliminação de itens para melhoria da precisão, especialmente se forem eliminados os itens que apresentaram cargas negativas neste fator. Porém, optou-se por realizar esse procedimento em um outro estudo posterior para evitar a eliminação somente por critério estatístico, quando também se deve levar em conta o conteúdo dos itens. Assim, basicamente, o primeiro fator secundário (FSO1) se relaciona à detecção de que não há informação suficiente para se concluir algo. Já o segundo fator secundário (FSO2) se relacionou à detecção de uma argumentação válida ou inválida.

A precisão desses fatores de segunda ordem foi de 0,844 para FSO1 e de 0,708 para o FSO2, que podem ser considerados bons tanto pelo SATEPSI quanto por critérios internacionais. Esses resultados sugerem que existe evidência de validade para uma estrutura bifatorial do teste de raciocínio lógico-dedutivo, relacionados a diferentes tipos de raciocínio, em que um deles permite concluir com segurança algo sobre a pergunta em questão (se é válido ou inválido) e outro que leva o respondente a concluir que não há informações suficientes para se concluir algo sobre o que é perguntado.

Diante de tal dedução da análise dos dados, procedeu-se a uma análise de correlação de Pearson entre os dois fatores secundários, encontrando-se um valor de 0,167, significativo ao nível de 0,022. Embora a correlação entre os dois fatores tenha sido positiva e estatisticamente significativa, ela foi muito baixa para dois conjuntos de itens que supostamente estariam avaliando o mesmo tipo de processamento cognitivo. Portanto, esse dado sugere que existem dois tipos de processamentos cognitivos subjacentes (representados pelo FSO1 e pelo FSO2) ao que se está chamando nessa prova de raciocínio lógico-dedutivo.

É possível que o impacto disso no cotidiano resulte em que as pessoas tendem a usar mais, por exemplo, o raciocínio para detectar se uma conclusão é verdadeira ou falsa, mas continuam procurando o verdadeiro ou falso mesmo quando não há informação suficiente para que uma conclusão seja afirmada dentro do conjunto de informações aos quais se tem acesso. Isso provavelmente está relacionado a dificuldades que ocorrem em decorrência das pessoas preencherem mentalmente informações inexistentes ou imprecisas nas premissas do teste em questão, o que as leva a tirar conclusões equivocadas, errando assim a resposta considerada correta pelo teste. O que seria interessante para desenvolver o pensamento crítico (tomado aqui com o conceito de reflexão ativa voltada para a resolução de um problema com máxima eficácia) seria estimular a passagem do FSO1 para o FSO2.

O terceiro objetivo deste estudo diz respeito a buscar evidências de validade baseadas na correlação com variáveis externas, como uma medida de raciocínio abstrato (validade convergente) e traços de personalidade dos cinco grandes fatores (validade discriminante). Para investigar a relação desses fatores com variáveis externas, foram realizadas análises de correlação de Pearson entre os fatores de segunda ordem do raciocínio lógico-dedutivo e os traços de personalidade e a prova de raciocínio abstrato. Os resultados estão na Tabela 3.

Tabela 3 - Correlação de Pearson dos fatores de segunda ordem obtidos no teste de raciocínio lógico-dedutivo com a prova de Raciocínio Abstrato (RA) da BPR5 e com os traços de personalidade da bateria BFP (Neuroticismo, Extroversão, Socialização, Responsabilidade e Abertura à experiência)

	FSO1	FSO2
RA total	0,406**	0,305**
N	-0,029	-0,120
E	-0,095	0,003
S	0,005	0,258*
R	0,004	0,220*
A	0,151	0,104

** . A correlação é significativa ao nível de 0.01 (2-caudas).

*. A correlação é significativa ao nível de 0.05 (2-caudas).

Os dados apresentados na Tabela 3 apoiam a validade convergente do instrumento, pois os dois fatores de raciocínio lógico-dedutivo se correlacionaram moderadamente, mas significativamente, com as pontuações no raciocínio abstrato mostrando que eles têm variância em comum ao mesmo tempo em que apresentam alguma variância única em relação a uma medida tradicional de inteligência. Isso seria esperado para instrumentos que avaliam tipos diferentes de raciocínio ao mesmo tempo em que apresentam algo em comum, provavelmente relacionado ao fator geral de inteligência.

Por sua vez a maioria dos valores de correlações com o teste de personalidade não foi significativa e apenas duas correlações foram significativas mas ainda assim com valores baixos. No conjunto, esses dados atestam a validade discriminante do instrumento, que se correlaciona mais fortemente com uma medida de inteligência do que com medidas de traços de personalidade.

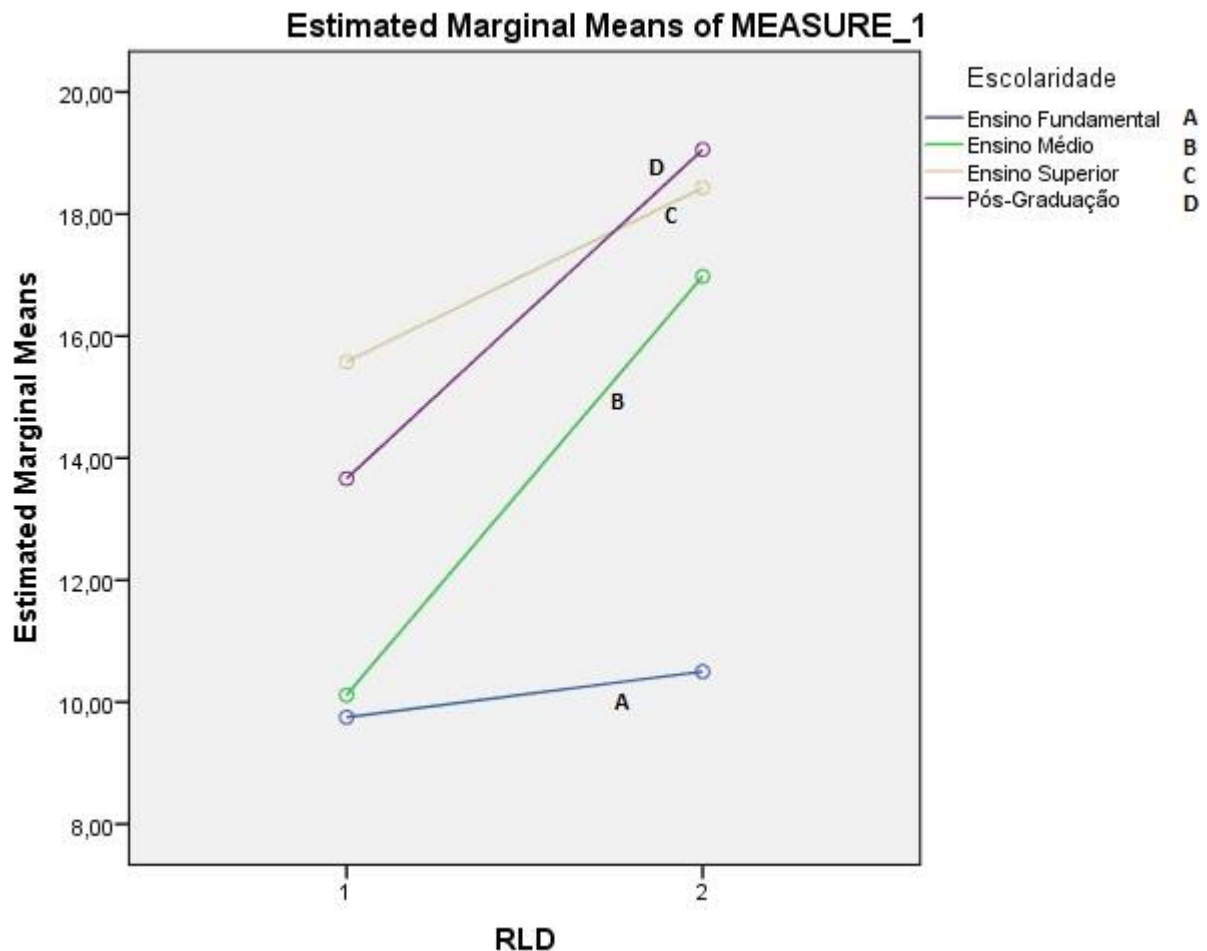
Outra análise correlacional foi realizada com os fatores de segunda ordem e alguns dados sócio-econômicos, produzindo a Tabela 4. Nesta tabela encontram-se informações sobre uma análise de variância de medidas repetidas para verificar o efeito das variáveis sexo, escolaridade e a da interação entre sexo e escolaridade nos Fatores de Segunda Ordem 1 e 2.

Tabela 4: Analise de Perfis de Medidas Repetidas (Tabachinick & Fidell, 1996)

Tests of Between-Subjects Effects						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	14385,901	1	14385,901	541,545	,000	,755
Sexo	7,797	1	7,797	,294	,589	,002
Escolaridade	1049,421	3	349,807	13,168	,000	,183
Sexo * Escolar	78,710	3	26,237	,988	,400	,017
Error	4675,358	176	26,565			

Foi observado que apenas a variável escolaridade mostrou-se estatisticamente significativa, sendo capaz de explicar 18,3% da variância total nas pontuações nos fatores de raciocínio lógico-dedutivo, conforme mostra a Figura 2 abaixo.

Figura 2: Correlação entre os Fatores de segunda ordem e o nível de escolaridade



Conforme se pode observar, as pontuações em ambos os fatores secundários obtidos (ponto 1 na linha horizontal (FSO1) e ponto 2 na linha horizontal (FSO2)) tendem a aumentar de acordo com a evolução da escolaridade. No entanto, o primeiro fator discrimina melhor a passagem do ensino fundamental (linha A) e médio (linha B) para o ensino superior (linha C) do que o segundo fator (em que o ensino médio aparece no mesmo nível do ensino superior), pois no FSO1 observa-se distanciamento dos valores obtidos pelo ensino fundamental e médio quando comparados com o valor obtido pelo ensino superior.

O resultado obtido nesta análise está de acordo com uma pesquisa realizada com universitários por Souza (2007), na qual foi encontrada correlação positiva entre um exame geral de rendimento acadêmico e o TRLD, revelando que ocorreu evolução do raciocínio lógico dedutivo no decorrer da vida acadêmica dos participantes da pesquisa. Assim, esses dados podem ser tomados como evidências de validade para o TRLD, pois a habilidade dos sujeitos nesse tipo de raciocínio aumenta conforme a escolaridade se eleva. Esse tipo de resultado é comumente encontrado em testes tradicionais de inteligência.

DISCUSSÃO DOS RESULTADOS

Este estudo pretendeu avaliar as propriedades psicométricas de um teste de raciocínio lógico-dedutivo utilizando participantes brasileiros. Como análise foi empregada a TCT a qual apontou resultados robustos, com formação de fatores primários e secundários com índices psicométricos significativos e válidos.

A análise fatorial permitiu a identificação de cinco fatores primários e dois fatores de segunda ordem, cujos itens se agruparam de acordo com características do tipo de resposta exigida no TRLD (verdadeira (v), falsa (f) ou não é possível ter certeza (p)). Dada a redundância de interpretações ocorrida nos fatores primários e a ocorrência de um fator com fidedignidade inadequada, considera-se mais apropriada a utilização da solução bifatorial de segunda ordem em pesquisas subsequentes. Em relação investigação da fidedignidade do instrumento, ambos os fatores de segunda ordem obtiveram índices acima de 0,7, atendendo tanto às normas do Sistema de Avaliação de Testes Psicológicos (SATEPSI) do Conselho Federal de Psicologia quanto às normas internacionais (Anastasi & Urbina, 2000; Pasquali, 1999, 2003; Urbina, 2007). Esses índices de precisão são bem superiores aos obtidos em pesquisas anteriores com outros instrumentos, como a de Nieto, Saiz e Orgaz (2009) utilizando o HCTAES e a de Rivas e Saiz (2012) com a PENCRISAL, citadas anteriormente.

Como terceiro objetivo da presente pesquisa, buscou-se evidências de validade baseadas na correlação com variáveis externas, através de análises de correlação de Pearson entre os fatores

de segunda ordem do raciocínio lógico-dedutivo e os traços de personalidade e a prova de raciocínio abstrato. Os dados encontrados nesta análise apoiaram a validade convergente do instrumento, uma vez que foi observado que os dois fatores de raciocínio lógico-dedutivo se correlacionaram significativamente e as pontuações no teste de raciocínio abstrato mostraram que existe variância em comum e também apresentam variância única em relação a uma medida tradicional de inteligência, ou seja, o teste de raciocínio lógico-dedutivo e o teste de raciocínio abstrato apresentaram algo em comum, que sugere relação com o fator geral de inteligência. Adicionalmente, o Teste de Raciocínio Lógico-Dedutivo apresentou um padrão de correlações baixas (a maioria não significativas) com traços de personalidade, o que atesta a sua validade discriminante. Por fim foi encontrada correlação significativa dos fatores de segunda ordem com o nível de escolaridade, o que aponta para evidência de validade desenvolvimental para o TRLD.

Assim, esta pesquisa contribui com a área de avaliação psicológica de funções cognitivas ao apresentar evidências de validade e precisão para um instrumento de avaliação do raciocínio lógico-dedutivo. Essas evidências, no entanto, devem ser tomadas como um indicativo de que vale a pena investir no desenvolvimento desse instrumento, mas não como palavra final a respeito de suas propriedades psicométricas. Os resultados deste estudo foram obtidos com uma amostra bastante específica, que em sua maioria eram jovens adultos, moradores da região metropolitana do Recife-PE, o que restringe a possibilidade de generalização dos resultados.

Além disso, os dados sugerem a necessidade de realização de outros estudos para o desenvolvimento do TRLD. Por exemplo, é possível que alguns itens possam ser eliminados sem perda de precisão, mas com redução da extensão do instrumento. Para isso, seria interessante analisar os dois fatores de segunda ordem com auxílio da Teoria de Resposta ao Item, por exemplo, que poderia trazer mais informações pormenorizadas sobre o funcionamento dos itens dentro de cada fator, auxiliando na escolha dos melhores itens para a composição de uma nova versão do instrumento. Apesar disso, essa pesquisa acrescenta conhecimento científico ao mostrar que existem dois aspectos relacionados ao processamento do raciocínio lógico-dedutivo e ao apresentar evidências de validade e fidedignidade para uma medida adequada desse construto. Assim, esse instrumento pode ser recomendado para uso em pesquisas.

REFERÊNCIAS

- Anastasi, A; Urbina, S. (2000). *Testagem Psicológica*. Porto Alegre: Artes Médicas. pp.10-95.
- Bernard, R.M.; Zhang, D.; Abrami, P. C.; Sicol, F.; Borokhovski, E.; Surkes, M. A. (2008). Exploring the structure of the Watson–Glaser Critical Thinking Appraisal: One scale or many subscales? *Thinking Skills and Creativity*, v.3(1), pp.15-22.
- Bernardes, M. E. M; (2011). O pensamento na atividade prática: implicações no processo pedagógico. *Psicologia em estudo*, Maringá, v. 16, n.4. pp.521-530, out/dez.
- Copi, Irving. (1974). *Introdução à lógica*. São Paulo: Ed Mestre Jou, pp.19-39.
- Ennis, R. H. (1996). *Critical thinking*. Upper Saddle River, NJ: Prentice-Hall.
- Gueroult, M. (2007). Lógica, arquitetura e estruturas constitutivas dos sistemas filosóficos. Tradução: Pedro Jonas de Almeida. *Trans/Form/Ação*, São Paulo, 30(1): pp.235-246.
- Haack, Susan; (2002). *Filosofia das lógicas*. Tradução: Cezar Augusto Mortari; Luiz Henrique de Araújo Dutra. São Paulo: Editora UNESP. pp. 25-51.
- Halpern, D. F. (2006). *Halpern Critical Thinking Assessment Using Everyday Situations: Background and scoring standards*. Unpublished report.
- Mundim, R. P. (2002). A lógica formal – princípios elementares. *Economia & gestão*. Belo Horizonte, v.2, n.3. pp.135-145.
- Muthen, L., & Muthen, B. (2010). Mplus 6.11 version user's guide. Los Angeles, CA.
- Nieto, A.M.; Saiz, C.; Orgaz, B. (2009). Análisis de la propiedades psicométricas de la versión española del HCTAES-Test de Halpern para la evaluación del pensamiento crítico mediante situaciones cotidianas. *Revista electrónica de metodología aplicada*, v.14, n.1, pp.1-15.
- Nolt, J & Rohatyn, D. (1991). *Lógica*. Tradução Mineko Yamashita. São Paulo: McGraw-Hill. pp. 01-65.
- Nunes, C.H.S.S; Huts, C.S; Nunes, M.F.O. (2010). Bateria Fatorial de Personalidade (BFP): *Manual Técnico*. São Paulo: Casa do psicólogo.
- Pasquali, L. (1999). Testes referentes a construto: Teoria e modelo de construção. In L. Pasquali (Ed.), *Instrumentos psicológicos: Manual prático de elaboração*. Brasília, DF: Laboratório de Pesquisa em Avaliação e Medida. pp. 56-75.
- Pasquali, L. (2003). *Psicometria: Teoria dos Testes na Psicologia e na Educação*. Petrópolis: Vozes. pp. 67-105.

- Pinto, A.V. (1969). *Ciência e existência: problemas filosóficos da pesquisa científica*. Rio de Janeiro: Paz e Terra. pp. 10-140.
- Primi, R.; Almeida, L.S. (2000). Estudo de validação da bateria de provas de raciocínio (BPR-5). *Psicologia: Teoria e pesquisa*. Vol.16, n.2, pp.165-173.
- Rivas, Silvia F. & Saiz, Carlos (2012). Validación y propiedades psicométricas de La prueba de pensamiento crítico – PENCRISAL. *Revista Electrónica de Metodología Aplicada*. Vol.17, n 1, pp.18-34.
- Saiz, Carlos. (2002) *Pensamiento crítico – Conceptos básicos y actividades prácticas*. Psicología pirámide, pp.16-28.
- Souza, M.A.V.F.; (2007). *Solução de problemas: relações entre habilidade matemática, representação mental, desempenho e raciocínio dedutivo*. Tese, UNICAMP, Campinas.
- Tabachinick, B. G. & Fidell, L. S. (1996). *Using multivariate statistics*. New York: HarperCollins.
- Urbina, S. (2007). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artmed. pp. 121-211.
- West, R.F.; Toplak, M.E & Stanovich, K.E., (2008). Heuristics and biases as measures of critical thinking: associations with cognitive abilities and thinking dispositions. *Journal of educational psychology*, 100(4), pp.930-941.

Amorim, M. (2015). Análise de um teste de raciocínio lógico-dedutivo com TRI. Dissertação de Mestrado, Pós-Graduação em Psicologia Cognitiva, Universidade Federal de Pernambuco, Recife.

RESUMO

O presente estudo analisou os dados estatísticos sobre o pensamento crítico numa amostra de 192 participantes brasileiros (66 do sexo masculino e 126 do sexo feminino), com idade superior a 18 anos, aos quais foram aplicados o teste de Raciocínio Lógico-Dedutivo (desenvolvido por Primi). O objetivo geral deste trabalho foi investigar as propriedades psicométricas por meio da Teoria de Reposta ao Item (TRI), no intuito de obter análise do ajuste dos dados ao Modelo de Rasch da TRI (*infit* e *outfit*); Analisar a adequação do nível de dificuldade dos itens (parâmetro b) ao nível de habilidade dos sujeitos θ (Θ) e por fim analisar o espectro de precisão do instrumento. Após análise fatorial foram extraídos dois Fatores de Segunda Ordem, o FSO1 e o FSO2, com precisão de 0,844 para o FSO1 e de 0,708 para o FSO2. Uma análise de correlação de Pearson resultou numa correlação estatisticamente significativa ($p=0,022$) de 0,167. Foi constatado que o FSO1 está relacionado ao raciocínio necessário para se concluir que existe pouca informação para se afirmar algo no item, enquanto o FSO2 está relacionado com a detecção de uma resposta, válida ou inválida, que se pode retirar do item. Em relação à análise com a TRI no FSO1 os índices de *infit* variaram de 0,82 a 1,41 ($M=1,00$; $DP=0,16$) sendo considerados indicativos de bom ajustamento pois se encontram dentro do intervalo de 0,5 e 1,5. Por sua vez os índices de *outfit* variaram de 0,73 a 1,77 ($M=1,02$; $DP=0,26$). Já a análise do FSO2 apresentou índices de *infit* variando de 0,83 a 1,42 ($M=0,99$; $DP=0,13$). Os índices de *outfit* variaram de 0,34 a 2,59 ($M=0,90$; $DP=0,50$). No geral, pode-se dizer que os dados se ajustaram ao modelo matemático da TRI.

Palavras-chave: TRI; pensamento crítico; avaliação psicológica; testes psicológicos.

Amorim, Mozart. (2015). Analysis of a logical-deductive reasoning test with TRI. Master's degree thesis, Cognitive Psychology Graduate School, Universidade Federal de Pernambuco, Recife.

ABSTRACT

This study analyzed the statistical data on critical thinking in a sample of 192 participants (66 male and 126 female) in northeastern Brazil (older than 18 years) to which were applied the logics with deductive (developed by Primi). The aim of this study was to investigate the psychometric properties by Item Response Theory (IRT) in order to obtain analysis fitting the data to the Rasch model IRT (infit and outfit); To analyze the adequacy of the level of difficulty of the items (parameter b) the skill level of theta subjects (θ) and finally analyze the instrument's accuracy spectrum. After factor analysis were extracted two factors of Second Order, the FSO1 and FSO2 accurate to 0.844 for FSO1 and 0.708 for FSO2. A Pearson correlation analysis resulted in a statistically significant correlation ($p = 0.022$) 0.167. It was found that the FSO1 is related to reasoning to conclude that there is little information to say something on the item, while the FSO2 associated with the detection of an answer, valid or invalid, that can be drawn from the item. As regards the analysis with the IRT FSO1 the infit rates varied from 0.82 to 1.41 (mean = 1.00; SD = 0.16) was considered indicative of a good fit as they are within the range of 0, 5 and 1.5. In turn the outfit rates varied from 0.73 to 1.77 ($M = 1.02$, $SD = 0.26$). The analysis of FSO2 presented infit rates ranging from 0.83 to 1.42 ($M = 0.99$, $SD = 0.13$). The outfit rates varied from 0.34 to 2.59 ($M = 0.90$, $SD = 0.50$). Overall, it can be said that the data fitted to the mathematical model of the IRT.

Keywords: IRT; critical thinking; psychological assessment; psychological tests.

INTRODUÇÃO

O ser humano atribui aos seus processos mentais a capacidade de observação e interpretação, abstraindo formas e dinâmicas do mundo real, obtendo conhecimentos diversos os quais possibilitam a sua interferência no mundo, nos objetos e também nos demais seres vivos. Contudo, não se faz necessário apenas observar e agir neste meio, pois sem o domínio do processo de raciocínio o ser humano basicamente pouco se distinguiria dos outros animais. A observação de nada serviria sem o atributo da razão, a qual demanda reflexão sobre determinado problema ou questão abstrata ou real, possibilitando ao ser humano superar o seu conhecimento atual diante de seu meio físico e social, e por sua vez abrindo-lhe novas possibilidades de captar novas realidades a partir da análise ou avaliação de si mesmo e de seu meio (Bernardes, 2011; Kopnin, 1978).

Articular informações para produzir uma conclusão racional sobre algo se constitui num processo argumentativo, o qual demanda processamentos cognitivos diversos que envolvem o raciocínio dedutivo, o raciocínio indutivo, a reflexão, entre outros, sempre presente nesta atitude humana de análise do mundo, dos outros e de si. Através do uso da razão o ser humano é capaz de refletir sobre informações do meio, retirando conclusões, procurando assim resolver problemas reais ou teóricos (Bernardes, 2011; Nolt & Rohatyn, 1991).

Em se tratando das formas de argumentação, existe a lógica, a qual estuda a validade ou não dos argumentos, utilizando o raciocínio indutivo (partindo de premissas particulares e chegando a uma conclusão universal) e também o raciocínio dedutivo (utilizando premissas universais para atingir uma conclusão particular). No que diz respeito aos argumentos estruturados em linguagem comum existe o ramo da lógica informal, a qual permite a análise de estruturas argumentativas dispostas em linguagem utilizada no cotidiano, mas que mesmo assim necessita passar pelo crivo da validação ou não da mensagem pelo receptor. Tal processo permite que o receptor indique se a mensagem argumentativa é válida ou não, ou seja, distinguir o raciocínio válido das falácias envolve pensar criticamente sobre a mensagem recebida (Nolt & Rohatyn, 1991; Copi, 1974,).

Dentro das estruturas de linguagem utilizada pela lógica, a semântica lógica é um tema que trata da construção de argumentos consistentes com base nas constantes semânticas (significados contínuos ou frequentes), uma vez que um requisito básico da argumentação crítica é

a de que um crítico deve tentar avaliar o contexto do diálogo em que o argumento ocorre. Isto implica por vezes que devemos lidar com argumentos longos e também complexos, obrigando desta forma uma investigação profunda de partes não declaradas da argumentação textual, conforme assinalado por algum tipo de evidência contida no texto, com o objetivo de que o argumento crítico teria que responder a uma dada pergunta. Tal fato leva a concluir que uma crítica deve ser feita a partir de uma discussão, permitindo classificar o argumento como fraco, errôneo ou falacioso, sempre levando-se em consideração de que a justificativa primordial para considerar razoável as críticas que foram lançadas, devem ser dada a partir de evidências documentais do próprio texto e também do contexto em que nos foi dado o argumento (Walton, 1989).

O pensamento crítico encontra-se inserido no campo da lógica informal, permitindo investigar outras formas plausíveis de raciocínio que não fossem somente as estruturas formais. Desta forma a análise e estudo do pensamento crítico se desenvolveu dentro deste ramo da lógica, tomando como cerne de sua observação a argumentação voltada para as habilidades fundamentais utilizadas para a comprovação de hipóteses, tomada de decisão e resolução de problemas. O pensamento crítico engloba tanto um ato cognitivo como um ato emocional, uma vez que as emoções atuam durante o processo de racionalização, e desta forma existe a necessidade de que o controle emocional seja mantido afim de facilitar seu desencadeamento satisfatório (Brookfield, 1987; Ennis, 1996; Saiz, 2002).

O conceito de pensamento crítico pode ser considerado vasto mas ao mesmo tempo passível de concordância entre os diversos estudiosos que se debruçam sobre o tema. Assim Ennis (1989) considera que pensamento crítico seria a ação de um sujeito pensante o qual consegue refletir moderadamente sobre o que acreditar e de como operar diante de alguma situação que lhe foi colocada.

Por sua vez McPeck (1981) propõe que o pensamento crítico pode ser conceituado como uma habilidade de engajar-se ceticamente numa atividade reflexiva/refletida, demandando profunda reflexão acerca de algo. Por fim temos uma outra definição que unifica as duas anteriores, proposta por Halonen (1995) a qual atribui ao pensamento crítico a habilidade de engajamento cético em atividade reflexiva/refletida, tomando como foco a decisão sobre o que acreditar e de como atuar diante de algo.

Segundo West, Toplak & Stanovich (2008) é possível considerar o pensamento crítico como um subdomínio do pensamento racional, no qual seria considerado uma competência usualmente utilizada em processos de tomada de decisão.

Ainda no que concerne ao conceito, Halpern (2003) fala a respeito de que o pensamento crítico permite o desenvolvimento de comportamentos ou práticas que visam atingir uma meta específica, a qual será possível após análise dos argumentos dados e a testagem de hipóteses que por fim permitirá algum tipo de tomada de decisão para alcance da meta. Tomando como cerne o final deste processo, Paul (2005) define o pensamento crítico como a capacidade de produção de boas decisões a partir do próprio processo de tomada de decisão.

Por sua vez, Brady (2008) conceitua o pensamento crítico a partir da caracterização de raciocínio complexo, possuindo qualidade superior e com capacidade de articulação entre o conhecimento, a experiência e as competências intelectuais do indivíduo.

No que diz respeito ao contexto educacional, o pensamento crítico se torna relevante pois permite o desenvolvimento de habilidades que tornam possível desafiar novas e distintas situações exigidas, como também permite o desenvolvimento de proposição de soluções ajustadas para cada tipo de enfrentamento experienciado, onde o ser humano perseguiria sua própria autonomia com iniciativa própria (Barnes, 2005; Carelli & Santos, 1998). Em se tratando mais especificamente do nível superior de ensino, esta capacidade de autorregulação do aluno é desejável, uma vez que ela permite evolução acadêmica e educacional constantes, possibilitando maior eficácia do aluno diante das situações complexas que lhe são colocadas no âmbito educacional bem como futuramente no âmbito profissional (Almeida; Guisande; Pereira; Joly; Donaciano; Mendes; Ribeiro, 2009). Por fim o desenvolvimento do pensamento crítico torna-se importante no contexto educacional por ter a capacidade de desenvolver os procesos relativos à tomada de decisão, bem como permitir adaptações desejáveis no âmbito educacional que dizem respeito aos aspectos cognitivos e as características comportamentais necessárias para uma adequada evolução escolar (Lindner & Harris, 1993; Saiz & Rivas, 2010; Pintrich, Smith, Garcia & McKeachie, 1993; Paul, 2005).

Sobre as formas de avaliação do pensamento crítico, existem alguns instrumentos desenvolvidos fora do Brasil que já possuem vários estudos sobre suas propriedades psicométricas. Um destes instrumentos é o *Halpern Critical Thinking Assessment Using Everyday Situations – HCTAES* (Halpern, 2006), que se propõe a avaliar o pensamento crítico através de problemáticas

relacionadas com situações cotidianas, exigindo respostas abertas e também do tipo fechadas das pessoas participantes. Esta avaliação desenvolveu-se a partir da concepção de que o pensamento crítico exija cinco aptidões: comprovação de hipóteses, raciocínio verbal, análise do argumento, probabilidade e incerteza e tomada de decisão (Halpern, 1998, 2010). Em pesquisa realizada por Nieto, Saiz e Orgaz (2009) foram encontrados Alfa de Cronbach com os valores de 0,635 para a habilidade de comprovação de hipóteses, 0,367 para a habilidade de raciocínio verbal, 0,465 para a análise do argumento, 0,340 para probabilidade e incerteza e por fim o valor de 0,487 para a habilidade de tomada de decisão. Vale ressaltar que estes índices encontrados não atingiram valores mínimos exigidos por parâmetros internacionais (Urbina, 2007). A pesquisa encontrou medida de adequação da amostra através da prova de Kaiser-Mayer-Olkin com valor de 0,70 e prova de esfericidade de Bartlett ($\chi^2=2214,44$, $p<0,05$), o que levou os pesquisadores a concluir que o instrumento avaliava uma única dimensão, ou seja, o pensamento crítico seria fundamentado por um conjunto de habilidades presentes nos fatores analisados na pesquisa.

Também existe um outro instrumental chamado *Watson-Glaser Critical Thinking Assessment*, que possui bateria com 80 itens exigindo resposta do tipo múltipla escolha e se propondo a avaliar cinco dimensões: inferências, reconhecimento de pressupostos, dedução, interpretação e avaliação de argumentos (West, Toplak & Stanovich, 2008).

Rivas e Saiz (2012) por sua vez desenvolveram um tipo de avaliação do pensamento crítico baseada em respostas do tipo abertas e fundamentada em três habilidades principais: o raciocínio, a resolução de problemas e a tomada de decisão, chamando seu instrumental de teste PENCRISAL. Este teste utiliza itens com problemáticas que abordam questões rotineiras, solicitando desta forma que o participante utilize diferentes competências mentais, como por exemplo o raciocínio prático, o raciocínio indutivo, a tomada de decisão, a resolução de problemas, e por fim o raciocínio dedutivo. Realizando uma pesquisa de validação do PENCRISAL Rivas e Saiz (2012) encontraram para a consistência interna dos 35 itens (Alfa de Cronbach) o valor de 0,632 com $p<0,001$. A precisão entre avaliadores obteve valores de Kappa dentro do intervalo de 0,6 a 0,9, com $p<0,001$. No que diz respeito às cargas fatoriais, para o fator dedução o alfa de Cronbach atingiu 0,371 (KMO=0,634; Bartlett $p<0,001$), respondendo assim por 10% da variância total do teste. Já o fator indução respondeu por 11% da variância do PENCRISAL, obtendo alfa de Cronbach igual a 0,250 (KMO=0,575; Bartlett $p<0,001$). Por sua vez o fator denominado raciocínio prático atingiu alfa de Cronbach de 0,425 (KMO=0,624; Bartlett $p<0,001$), explicando 9% da

variância do teste. Em seguida o fator tomada de decisão alcançou alfa de Cronbach de 0,213 (KMO=0,575; Bartlett $p<0,001$) e respondeu por 15% da variância total. Finalizando, o fator Resolução de problemas chegou ao alfa de Cronbach no valor de 0,373 (KMO= 0,624; Bartlett $p<0,001$), explicando assim 8,4% da variância total do PENCRISAL. Esses valores foram considerados como bons pelos autores (Rivas & Saiz, 2012).

No Brasil existe um instrumento para avaliação do raciocínio lógico-dedutivo (Primi & Muller, não publicado), o qual foi utilizado em estudo de Amorim e Bueno (manuscrito não submetido) e essa pesquisa pretendeu avaliar suas propriedades psicométricas pela Teoria Clássica dos Testes. Foram encontrados cinco fatores primários, com coeficiente de Kuder-Richardson superiores a 0,6, com exceção do Fator 5, que ficou com 0,578. De forma geral, os fatores 1 e 5 se referiam à capacidade de identificar se uma conclusão era verdadeira ou falsa, enquanto os fatores 2, 3 e 4 se referiam à habilidade de detectar que não havia informações suficientes para se chegar a alguma conclusão confiável.

A possibilidade desse agrupamento foi confirmada por uma análise fatorial de segunda ordem, que resultou numa solução bifatorial, com os fatores primários 1 e 5 numa dimensão (FSO2) e os fatores primários 2, 3 e 4 em outra dimensão secundária (FSO1). A precisão desses fatores foi de 0,844 e de 0,708 para os dois fatores de segunda ordem, respectivamente. Uma análise de correlação de Pearson mostrou que esses fatores estavam pobremente correlacionados entre si, com uma correlação de 0,167 ($p=0,022$). Apesar disso, ambos os fatores se relacionaram moderadamente com uma medida de Raciocínio Abstrato e pobremente com medidas de traços de personalidade. Esses dados foram interpretados como promissores, mas os autores sugerem a continuidade no desenvolvimento do instrumento, recomendando que uma análise com TRI poderia ser útil na identificação de itens que pudessem ser eliminados sem comprometimento da precisão e da distribuição da dificuldade dos itens por toda a extensão de habilidade dos sujeitos.

Por isso, o objetivo geral deste trabalho foi investigar as propriedades psicométricas por meio da Teoria de Reposta ao Item (TRI), no intuito de obter análise do ajuste dos dados ao Modelo de Rasch da TRI (*infit* e *outfit*); analisar a adequação do nível de dificuldade dos itens (parâmetro b) ao nível de habilidade dos sujeitos θ (Θ) e, por fim, analisar a adequação das respostas consideradas corretas.

MÉTODO

Participantes

A amostra do estudo foi composta por um quantitativo de 192 participantes (66 do sexo masculino e 126 do sexo feminino), todos com idade superior a 18 anos, sendo a idade inferior a 18 anos o critério de exclusão. Não foi exigido qualquer nível de escolaridade.

Instrumentos

Inicialmente a pesquisa coletou dados sócio-econômicos, através de um questionário de informações pessoais, preenchidos por cada participante antes da aplicação do teste. Nesta pesquisa foi utilizado o Teste de Raciocínio Lógico-Dedutivo - TRLD (Primi & Muller, não publicado), pretendendo avaliar a capacidade de analisar informações, articulando-as no intuito de solucionar problemas presentes nos itens do teste. Quarenta e seis itens e subitens compõem o TRLD e sua pontuação é do tipo dicotômica (certo ou errado). O escore máximo alcançado pelo participante é de quarenta e seis enquanto o valor mínimo possível é de zero.

Procedimentos

Após submissão e aprovação do projeto no comitê de ética para pesquisas com seres humanos da Universidade Federal de Pernambuco foi iniciada a etapa de coleta de dados deste estudo. Foram repassados aos participantes todas as informações acerca dos objetivos bem como sobre os procedimentos envolvidos neste estudo. Tais informações também estavam acessíveis no Termo de Consentimento Livre e Esclarecido (TCLE), o qual constitui-se numa exigência do

próprio comitê de ética, e neste foi dado tempo para que todos os participantes da pesquisa lessem o TCLE e retirassem suas possíveis dúvidas. Após concordarem com o TCLE e assinarem, os participantes permitiram a utilização dos dados coletados de suas respostas no instrumento para utilização científica com objetivos de análises estatísticas, respeitando assim todas as demais informações pessoais fornecidas na fase de coleta.

A cada participante foi explanado que existiria o benefício direto de receber os dados de sua performance ao final completo da pesquisa. Como benefício indireto o participante ficou ciente de que sua contribuição em responder ao teste estaria contribuindo para a produção de conhecimento científico que poderia auxiliar na melhoria dos processos educacionais bem como em pesquisas futuras.

Todos os dados foram coletados por meio de instrumental impresso e com preenchimento de respostas em uma folha específica. Os participantes foram contatados na Universidade Federal de Pernambuco, mas também foram utilizados serviços eletrônicos como envio de email, enviando convite para participação neste estudo. Ao todo participaram da pesquisa 249 voluntários, mas pelos critérios de exclusão adotados, que consistiam na idade inferior a 18 anos, foi necessário proceder à exclusão dos dados de 57 participantes.

Finalizando a etapa de coleta de dados, o estudo ficou contemplado com o total de 192 participantes, onde todos responderam o TRI. As respostas de cada participante foram posteriormente analisadas de acordo com o padrão de respostas já existente para o teste, sendo então pontuadas e posteriormente digitalizadas em planilha eletrônica para prosseguir com a análise estatística.

A análise de dados foi realizada para os dois fatores de segunda ordem (FSO1 e FSO2), com auxílio do *software* Winsteps versão 3.69.1.6, e pretendeu verificar os seguintes pontos:

- a) O ajustamento dos itens ao modelo da TRI, utilizando-se os índices de *infit* e *outfit*. Segundo Linacre (2002) o índice *infit* demonstra o grau de ajustamento dos itens à região de habilidade em que o item é mais discriminativo. Já o índice de *outfit* diz respeito ao grau de ajustamento dos itens a outras regiões de habilidade. Índices com valores entre 0,5 e 1,5 são indicativos de um bom ajustamento, enquanto índices com valores entre 1,5 e 2,0 são considerados improdutivos, mas não degradantes para a medida. Por fim, valores acima de 2,0 são considerados degradantes para a medida.

- b) O nível de adequação do nível de dificuldade dos itens ao nível de habilidade dos participantes, que foi investigado, para cada fator secundário, por meio da comparação entre as médias das habilidades dos sujeitos (Θ) com a média da dificuldade dos itens (centradas em zero, por padrão do *software* Winsteps) e também pela inspeção do mapa de itens. Esperava-se que não houvesse discrepância entre essas médias e que os itens se distribuíssem ao longo do contínuo de habilidades, de modo que houvesse itens bons para discriminar sujeitos com baixas, médias e altas habilidades em ambos os fatores.
- c) O nível de adequação das alternativas de resposta consideradas corretas. Se as alternativas de resposta que foram consideradas corretas em cada item forem realmente as mais adequadas, elas devem ser escolhidas pelos sujeitos com maior média de habilidade (*theta*)

RESULTADOS

Os dados serão apresentados por fator e na ordem da proposição dos objetivos. Assim, utilizando-se análise de dados do FSO1, descrito anteriormente, analisou-se o ajustamento dos dados ao modelo e em seguida a adequação da dificuldade dos itens ao nível de habilidade dos sujeitos. Finalmente procedeu-se a adequação das alternativas de resposta consideradas corretas em cada item.

Em relação ao FSO1, pode-se observar que os índices de *infit* variaram de 0,82 a 1,41 ($M=1,00$; $DP=0,16$), sendo considerados indicativos de bom ajustamento, segundo Linacre (2002). Por sua vez os índices de *outfit* variaram de 0,73 a 1,77 ($M=1,02$; $DP=0,26$). Somente um item (i10) apresentou problema de ajustamento, obtendo *outfit* igual a 1,77, sendo considerado improdutivo, mas não degradante para a medida (Linacre, 2002). Problemas de ajustamento de *infit* são considerados de maior gravidade, já que estão relacionados a desajustes na região de habilidade onde o item é mais discriminativo, ou seja, onde se esperaria que ele funcionasse melhor. Problemas de ajustamento nessa região podem estar relacionadas com a dimensionalidade. Talvez o item não esteja tão associado ao construto, o que levaria a erro na plausibilidade de acerto do item pelos sujeitos. Já elevados índices de *outfit* indicam que problemas de ajustamento dos itens a regiões mais extremas de habilidade, para as quais o teste deve apresentar outros itens que sejam mais apropriados (discriminativos). Portanto, de modo geral, pode-se concluir que os dados estão ajustados ao modelo matemático da TRI.

A partir disso, para verificar a relação entre a dificuldade dos itens e a habilidade dos sujeitos, foi formado o mapa de itens do FSO1. A Figura 1 apresenta o mapa de itens, com a média de distribuição dos itens do lado direito, enquanto a média das habilidades dos sujeitos (Θ) se encontra do lado esquerdo.

questão o nível de dificuldade dos itens se mostrou elevado. Pelo mapa pode-se verificar que em se tratando de nível de dificuldade, tivemos o item i9 classificado como o mais difícil enquanto o item i31 foi classificado como o mais fácil.

Também é possível notar que conforme se distancia do ponto central (*logits* acima de 1 e abaixo de -1), especialmente na região de baixas habilidades, existem poucos itens para muitos participantes nesta região, ou seja, a dificuldade destes poucos itens encontra-se adaptada para uma grande quantidade de participantes, o que pode sugerir a necessidade de criação de novos itens, visando apenas melhorar a discriminação entre sujeitos, que para esta amostra se mostraram com baixas habilidades.

Por outro lado, foram encontrados conjuntos de itens (i16 e i33; i3, i37 e i41; i21, i32 e i36; i35 e i42; i22 e i45; i29 e i7; i19 e i27; i10 e i5) com nível de dificuldade considerados muito próximos, que talvez sugira a possibilidade de eliminação sem prejuízo para as propriedades de medida do instrumento, permitindo assim possíveis ganhos em termos de tempo e menor esforço para responder o teste.

O terceiro objetivo deste trabalho diz respeito à adequação das alternativas de resposta escolhidas como corretas. Caso a alternativa escolhida como correta esteja realmente adequada, ela deve ser escolhida por um grupo de sujeitos com média de habilidade superior á dos grupos de sujeitos que escolheram as outras alternativas (consideradas incorretas). Exemplos desse procedimento podem ser observados na Tabela 1.

Tabela 1 – Estatística descritiva dos itens do FSO1

Itens do teste	Nº de alternativas escolhida pelos participantes	Pontuação recebida (1 ou 0)	Número de participantes que escolheram a alternativa	Porcentagem de participantes que escolheram a alternativa	Média de theta (habilidade) dos participantes	S.E Mean	Outf	Correlação do item com o total do teste
9	1	0	2	1	-0,93	0,20	0,3	-0,7
	3	0	149	80	-0,34	0,09	1,1	-0,29
	2	1	35	19	0,63	0,26	1,5	0,32
			6	3	-0,62	0,28		-0,05
31	2	0	4	2	-1,71	0,22	0,3	-0,19
	1	0	8	4	-1,28	0,20	0,5	-0,20
	5	0	11	6	-1,10	0,20	0,6	-,020
	4	0	29	16	-0,91	0,15	0,9	-0,27
	3	1	134	72	0,19	0,10	0,9	0,47
			6	3	-0,94	1,10		-0,08

A Tabela 1 apresenta os dados relacionados às alternativas de resposta de apenas dois itens, como exemplo da análise realizada: um que foi considerado como o mais difícil (9) e outro que foi considerado como mais fácil (31). As colunas apresentam informações, respectivamente, sobre o número do item, o número da alternativa de resposta escolhida pelo participante, a pontuação que era dada para escolha de cada alternativa (a alternativa correta recebe um ponto, as demais zero), o número de pessoas que escolheram cada alternativa, a porcentagem de pessoas que escolheu cada alternativa, a média de habilidade (theta) das pessoas que escolheram cada alternativa.

No item 9, nota-se que a alternativa correta (2) não foi escolhida pelo maior número de participantes, mas as pessoas que escolheram essa alternativa apresentaram média de theta mais elevada em relação às outras alternativas. Outros itens se comportaram de forma semelhante ao item 9, ou seja, a alternativa correta não foi escolhida pela maioria dos participantes mas foi escolhida pelas pessoas com média de theta mais elevada. Esses itens foram: 5, 7, 10, 11, 19, 20 e 29.

Fato semelhante foi observado no item 31. Porém, nesse caso, a alternativa correta foi escolhida pela maioria das pessoas e essas também apresentaram média de theta superior aos que

escolheram outras alternativas. Outros itens se comportaram de forma semelhante ao i31, a saber: 1, 3, 16, 21, 22, 23, 25, 32, 33, 34, 35, 36, 37, 41, 42 e 45.

É importante destacar que em todos os itens desse fator, a alternativa considerada correta foi escolhida pelo grupo com média mais elevada de theta, indicando que a construção dos itens (especialmente, das alternativas consideradas corretas), foi bem-sucedida no que se refere a exigir maior habilidade para a escolha das alternativas corretas.

Procedeu-se da mesma forma para a análise de dados do FSO2 (analisou-se o ajustamento dos dados ao modelo, em seguida foi realizada a adequação da dificuldade dos itens ao nível de habilidade dos sujeitos e por fim a adequação das alternativas de resposta consideradas corretas em cada item). Esta análise do FSO2 apresentou índices de *infit* variando de 0,83 a 1,42 ($M=0,99$; $DP=0,13$), ou seja, dentro do intervalo desejável situado entre 0,5 e 1,5 (Linacre, 2002). Já os índices de *outfit* variaram de 0,34 a 2,59 ($M=0,90$; $DP=0,50$), demonstrando assim que um índice de *outfit* ficou fora do padrão com índice acima de 1,5 (i28), ou seja, este item apresentou índice que ultrapassa a aceitação máxima de interferência em um teste, pois indica que existe no item algum problema de ajustamento dele nas regiões mais extremas de habilidade, ou seja, ao que parece o i28 é pouco discriminativo. No geral, pode-se chegar à conclusão de que os dados se ajustam ao modelo da TRI (Linacre, 2002).

Após esta análise de *infit* e *outfit* procedeu-se à formação do mapa de itens do FSO2 (figura 2), permitindo melhor visualização entre a dificuldade dos itens (lado direito do mapa) e a habilidade dos participantes (lado esquerdo do mapa).

desenvolver outros itens com nível de dificuldade mais elevado, favorecendo uma melhor distribuição de itens por dificuldade, que deverão surgir no mapa em futuras análises, buscando uma forma de equilibrar melhor a dificuldade dos itens ao nível de habilidade dos sujeitos. Pelo mapa observa-se que o item mais difícil foi o i28, enquanto o mais fácil foi o i13.

Foi observado também que existem itens que devem estar avaliando o mesmo tipo de habilidade, sendo eles: i18 e i43; i24 e i6; i15 e i2, o que sugere a possibilidade de exclusão de alguns itens do teste, mais uma vez no intuito de permitir que ocorra menor desgaste do participante durante sua aplicação bem como a inserção de outros itens que possuam dificuldade mais elevada e que possam possibilitar um preenchimento melhor do mapa de itens bem como resultar num melhor equilíbrio entre os fatores secundários.

Para atingir o terceiro objetivo deste trabalho, que diz respeito à adequação das alternativas de resposta escolhidas como corretas, foi realizado o mesmo procedimento empregado para o FSO1. Exemplos desse procedimento para o FSO2 podem ser observados na Tabela 2.

Tabela 2 – Estatística descritiva dos itens do FSO2

Itens do teste	Nº de alternativas escolhida pelos participantes	Pontuação recebida (1 ou 0)	Número de participantes que escolheram a alternativa	Porcentagem de participantes que escolheram a alternativa	Média de theta (habilidade) dos participantes	S.E Mean	Outf	Correlação do item com o total do teste
28	3	0	29	16	1,25	0,23	0,7	-0,26
	1	0	98	53	2,10	0,10	1,5	0,08
	2	1	59	32	2,22	0,21	3,1	0,12
			6	3	-0,49	1,48		-0,15
13	2	0	2	1	-0,53	0,50	0,2	-0,21
	1	0	3	2	0,07	0,55	0,4	-0,20
	3	1	181	97	2,07	0,09	1	-0,29
			6	3	-0,44	1,44		-0,15

Conforme já realizado no FSO1, a Tabela 2 referente ao FSO2 apresenta os dados relacionados às alternativas de resposta de apenas dois itens, como exemplo da análise realizada:

um que foi considerado como o mais difícil (28) e outro que foi considerado como mais fácil (13). As colunas apresentam informações, respectivamente, sobre o número do item, o número da alternativa de resposta escolhida pelo participante, a pontuação que era dada para escolha de cada alternativa (a alternativa correta recebe um ponto, as demais zero), o número de pessoas que escolheram cada alternativa, a porcentagem de pessoas que escolheu cada alternativa, a média de habilidade (theta) das pessoas que escolheram cada alternativa.

No item 28 a alternativa correta (2) não foi escolhida pelo maior número de participantes, mas as pessoas que escolheram essa alternativa apresentaram média de theta mais elevada em relação às outras alternativas. O item i7 se comportou de forma semelhante ao item 28 ou seja, a alternativa correta não foi escolhida pela maioria dos participantes mas foi escolhida pelas pessoas com média de theta mais elevada.

Já em relação ao item 13, nota-se que a grande maioria (97% dos participantes) escolheu a alternativa correta e que também nesse caso estes participantes apresentaram média de theta superior aos que escolheram outras alternativas. Outros itens se comportaram de forma semelhante ao i13 foram: i2, i4, i6, i8, i14, i15, i17, i18, i24, i26, i39, i43, i44 e i46.

DISCUSSÃO DOS RESULTADOS

A pesquisa em tela avaliou as propriedades psicométricas do teste de raciocínio lógico-dedutivo através da formação de novos fatores de segunda ordem utilizando análise com a Teoria de Resposta ao Item (TRI). Os dados demonstraram que os índices de *infit* alcançaram níveis desejados, enquanto os índices de *outfit* ficaram dentro de um intervalo aceitável, contudo um índice ficou fora do padrão, apresentando valor acima de 1,5, o que sugere pouca discriminação do item em questão. Assim, exceto pelo último índice relatado, foi alcançado um dos objetivos desta pesquisa, pois os demais índices tanto de *infit* quanto de *outfit* ficaram no intervalo compreendido como bom ajustamento.

O segundo objetivo da pesquisa pretendeu analisar o nível de adequação, o nível de dificuldade dos itens e o nível de habilidade dos participantes, onde os mapas de itens demonstraram que, quando analisados simultaneamente, pode-se observar que a mesma amostra

de participantes encontrou maior dificuldade nos itens que compõem o FSO1 do que nos itens que formam o FSO2, uma vez que a média de *theta* do FSO1 ficou muito próxima da média dos itens. Tal dado sugere que o teste precisa ser reequilibrado contemplando mais itens que possuam menor dificuldade ao FSO1 e itens com maior dificuldade ao FSO2.

O terceiro objetivo da pesquisa pretendeu analisar o nível de adequação das alternativas de resposta consideradas corretas, onde era esperado que as alternativas de resposta correta em cada item fossem escolhidas pelos sujeitos com maior média de habilidade (*theta*). A análise constatou que tanto no FSO1 como no FSO2 os dois itens posicionados nas extremidades do mapa de itens, considerados itens mais fáceis (extremidade abaixo do zero) e os mais difíceis (extremidade acima do zero) apresentaram resultados semelhantes, pois as alternativas corretas foram selecionadas pelos participantes que apresentaram média de *theta* mais elevada em relação aos que escolheram outras alternativas.

Pode-se afirmar que a análise utilizando TRI forneceu contribuição acerca dos dados encontrados no TRLD, pois permitiu sugerir a exclusão de alguns itens, bem como a inserção de outros itens com novas características, permitindo assim discutir o instrumento em análise e produzir conhecimento científico na área. Todavia a amostra utilizada nesta pesquisa apresenta os vieses de ser predominantemente de moradores da região metropolitana do Recife-PE e por jovens adultos, restringindo a possibilidade de generalização dos resultados para pessoas com essas características. No entanto, os dados apresentados contribuem significativamente para o desenvolvimento psicométrico do instrumento, assim como serve de referência para o desenvolvimento de outros instrumentos de avaliação psicológica.

REFERÊNCIAS

- Almeida, L.S; Guisande, M.A; Pereira, A; Joly, M.C.R.A; Donaciano, B; Mendes, T; Ribeiro, M.S. (2009). Escala de competências de estudo (ECE-Sup): fundamentos e construção. In: B. D. Silva, L. S. Almeida, A. Barca. X. M. Peralbo (Orgs), *Atas do X Congresso Internacional Galego-Português de Psicopedagogia* (pp.4282-4292). Braga: CIED – Universidade do Minho.
- Barnes, C.A. (2005). Critical thinking revisited: Its past, present, and future. *New Directions for Community Colleges*, Summer 2005, pp. 5-13.
- Bernardes, M. E. M; (2011). O pensamento na atividade prática: implicações no processo pedagógico. *Psicologia em estudo*, Maringá, v. 16, n.4. pp.521-530, out/dez.
- Brady, M. (2008). Cover the material: or teach students to think? *Educational Leadership*, 65, pp.64-67.
- Brookfield, S.D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting*. San Francisco: Jossey-Bass.
- Carelli, M.J.G; Santos, A.A.A. (1998). Condições temporais e pessoais de estudo em universitários. *Psicologia escolar e educacional*, 2, 265-278.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and need research. *Educational Researcher*, 18, pp.13-16.
- Ennis, R. H. (1996). *Critical thinking*. Upper Saddle River, NJ: Prentice-Hall.
- Halonen, J. S. (1995). Desmystifying critical thinking. *Teaching of Psychology*, 22:1, pp.75-81.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American psychologist*, 53(4), pp.449-455.
- Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking (4th ed.)*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Halpern, D. F. (2006). *Halpern Critical Thinking Assessment Using Everyday Situations: Background and scoring standards*. Unpublished report.
- Halpern, D. F. (2010). *Manual HCTA, Halpern Critical Thinking Assessment (Version 2.1)*. Mödling, Áustria: Schuhfried.
- Kopnin, P.V. (1978). *A dialética como lógica e teoria do conhecimento*. Tradução: Paulo Bezerra. Rio de Janeiro: Civilização Brasileira.

- Linacre, J. M. (2002). What do Infit and Outfit, meansquare and standardized mean? *Rasch Measurement Transactions*, 16, 878.
- Lindner, R. W; Harris, B. (1993). Self-regulated learning: Its assessment and instructional implications. *Educational Research Quarterly*, 16, 29-37.
- McPeck, J.E. (1981). *Critical thinking and educations*. New York: St. Martins. pp. 05-20.
- Nieto, A.M.; Saiz, C.; Orgaz, B. (2009). Análisis de la propiedades psicométricas de la versión española del HCTAES-Test de Halpern para la evaluación del pensamiento crítico mediante situaciones cotidianas. *Revista electrónica de metodología aplicada*, v.14, n.1, pp.1-15.
- Nolt, J & Rohatyn, D. (1991). *Lógica*. Tradução Mineko Yamashita. São Paulo: McGraw-Hill. pp.01-65.
- Paul, R. (2005). The state of critical thinking today. *New directions for community colleges*, summer 2005, 27-38.
- Pintrich, P.R; Smith, D.A; Garcia, T; Mckeachie, W.J. (1993). Reability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement*, 53, 801-813.
- Rivas, Silvia F. & Saiz, Carlos (2012). Validación y propiedades psicométricas de La prueba de pensamiento crítico – PENCRISAL. *Revista Electrónica de Metodologia Aplicada*. v.17, n 1, pp.18-34.
- Saiz, Carlos. (2002) *Pensamiento crítico – Conceptos básicos y actividades prácticas*. Psicología pirámide, pp.16-28.
- Saiz, C. & Rivas, S. F. (2010). Melhorar El pensamiento critico contribuye al desarrollo personal de los jóvenes? In: H. J. Ribeiro & J. N. Vicente (Eds.), *O lugar da lógica e da argumentação no ensino da filosofia*. Coimbra: Unidade I & D, Linguagem, Interpretação e Filosofia. (pp. 39-52).
- Walton, D. N. (1989). *Informal Logic: a handbook for critical argumentation*. New York: Cambridge University Press. pp. 1-16.
- West, R.F.; Toplak, M.E & Stanovich, K.E., (2008). Heuristics and biases as measures of critical thinking: associations with cognitive abilities and thinking dispositions. *journal of educational psychology*, 100(4), pp. 930-941.

CONCLUSÃO GERAL

Os achados desta pesquisa forneceram bons resultados conforme os critérios estatísticos mínimos estabelecidos pelo SATEPSI. A primeira análise utilizando a teoria clássica de medidas forneceu resultados significativos, mas não encontrou um único fator no TRLD, como era esperado. Porém, foram encontrados cinco fatores principais e dois de segunda ordem, cuja interpretação foi relacionada ao tipo de resposta exigida para responder a situação-problema.

A análise dos fatores primários forneceu evidências confiáveis que apontaram para a possível necessidade de revisão de alguns itens que apresentaram cargas negativas. Esse fato foi totalmente inesperado uma vez que todos os itens eram supostos a medir o mesmo traço latente (raciocínio lógico-dedutivo). Por isso, os cinco fatores primários passaram por uma nova análise fatorial, que resultou em dois fatores de segunda ordem. Embora esses fatores tenham apresentado correlação baixa entre si, os fatores puderam ser associados a tipos de funcionamento cognitivo distintos, constituindo-se em evidências de validade com base na estrutura interna do instrumento.

Esses fatores se correlacionaram significativamente com uma medida de Raciocínio Abstrato, mas pobremente (predominância de correlações não-significativas) com traços de personalidade. Esses resultados se constituem em evidências de validade convergente e discriminante, respectivamente. Além disso, uma análise de variância multivariada, mostrou que a pontuação média nos fatores sobe concomitantemente ao aumento da escolaridade, fato que pode ser tomado como evidência de validade desenvolvimental.

Esse conjunto de resultados, somados aos bons índices de precisão obtidos, especialmente para os fatores de segunda ordem, sugerem que o instrumento apresenta bons indicadores de validade para avaliação de dois aspectos do raciocínio lógico-dedutivo: a capacidade em detectar se uma conclusão é verdadeira ou falsa (FSO1) e a capacidade em detectar que não há informações suficientes para se chegar a uma conclusão lógica. Não se conseguiu encontrar algo semelhante em estudos anteriores, podendo-se considerar esses resultados como uma contribuição importante deste trabalho para a área de avaliação de habilidades cognitivas. O segundo estudo, tendo os dados sido adequados ao modelo da TRI, constatou-se que há lacunas no contínuo de habilidades dos sujeitos que necessitam que novos itens sejam construídos para preenche-las. Por outro lado, há itens que avaliam o mesmo nível de habilidade, que poderiam ser eliminados, sem perda de

informação e com redução da extensão do instrumento. Esses dados sugerem que a continuidade do desenvolvimento do instrumento é necessária e pode levar uma melhoria na relação custo (tempo) – benefício. Além disso, as análises mostraram que as respostas escolhidas como corretas, foram todas escolhidas por um grupo de sujeitos que apresentou maior média de habilidade, mostrando a adequação dos itens para a discriminação entre sujeitos com diferentes habilidades.

Como limitações desse conjunto de estudos pode-se citar que a amostra foi composta predominantemente por participantes jovens adultos, moradores da região metropolitana do Recife. Diante de tal limitação, sugere-se que outros estudos possam ser realizados para medir as propriedades psicométricas do TRLD utilizando uma amostra mais ampla, tanto em termos numéricos como também em termos sócio-econômicos, permitindo assim verificar a estabilidade (ou não) dos resultados ora encontrados. De qualquer forma, o presente estudo forneceu dados estatísticos sustentáveis e com boas propriedades psicométricas, de acordo com níveis mínimos exigidos pelo SATEPSI do Conselho Federal de Psicologia, de modo que o Teste de Raciocínio Lógico Dedutivo, pode ser recomendado para uso em pesquisas.

ANEXO A

GfRLD – Teste de Raciocínio Lógico Dedutivo - FORMA A

Ricardo Primi & Franz Muller
Universidade São Francisco

Identificação _____

Nome: _____

INSTRUÇÕES

Nesta prova você terá a oportunidade de demonstrar sua capacidade de raciocinar de forma lógica para resolver problemas que lhe serão propostos. Leia atentamente cada questão e verifique o que é perguntado.

Bom trabalho!

Considere cada uma das 4 situações abaixo como a única fonte de verdade que você dispõe. Para cada situação apresentada há 4 conclusões. Julgue cada uma das conclusões usando:

- V necessariamente verdadeiro
- P não é possível ter certeza
- F necessariamente falso

1ª situação: Seu João organizou um passeio e todos os seus sobrinhos foram ao passeio. Você pode concluir que (assinale V, P ou F em cada uma das frases):

- a. V P F Todos que foram ao passeio são sobrinhos do seu João.
- b. V P F Alguns que foram ao passeio são sobrinhos do seu João.
- c. V P F Alguém que foi ao passeio não é sobrinho do seu João.
- d. V P F Ninguém que foi ao passeio é sobrinho do seu João.

2ª situação: Alguns cães usam coleira.

Você pode concluir que:

- a. V P F Todos que usam coleira são cães.
- b. V P F Alguns animais que usam coleira são cães.
- c. V P F Alguns animais que usam coleira não são cães.
- d. V P F Ninguém que usa coleira é cão.

3ª situação: Algumas plantas não são verdes.

Você pode concluir que:

- a. V P F Tudo que é verde é planta.
- b. V P F Algumas coisas verdes são plantas.
- c. V P F Algumas coisas verdes não são plantas.
- d. V P F Nada que é verde é planta.

4ª situação: Nenhuma cobra voa.

Você pode concluir que:

- a. V P F Tudo que voa é cobra.
- b. V P F Algumas coisas que voam são cobras.
- c. V P F Algumas coisas que voam não são cobras.
- d. V P F Nada que voa é cobra.

5. Em uma escola com 20 alunos são ensinados os idiomas inglês, francês e alemão. Todos os que estudam inglês estudam também francês. Nenhum aluno que estuda francês estuda alemão. Considerando que 13 alunos não estudam alemão, qual das conclusões abaixo **é verdadeira**?

- a. 13 alunos não estudam inglês.
- b. Apenas 7 alunos estudam inglês.
- c. 13 alunos estudam francês.
- d. Mais de 13 alunos estudam inglês.
- e. Não sei.

6. Uma loja vende produtos importados e nacionais, entre vestidos, camisas e casacos. Alguns vestidos e todos os casacos fazem parte dos produtos importados. Não há produto importado disponível em tamanho grande. Assinale qual dentre os fatos enunciados **NÃO** poderia ser verdadeiro:

- a. Carla experimenta uma camisa nacional.
- b. Luciana está comprando um casaco pequeno.
- c. Alberto pegou um casaco grande.
- d. Adriana experimenta um vestido pequeno.
- e. Fernanda experimenta um vestido grande.

7. Nenhum aluno da professora Ana estudou filosofia.

Todos os alunos que estudaram filosofia foram ao churrasco.

Responda cada um dos itens abaixo usando:

- V necessariamente verdadeiro
- P não é possível ter certeza
- F necessariamente falso

- a. V P F Alguns que foram ao churrasco são alunos da professora Ana.
- b. V P F Nenhum aluno da professora Ana foi ao churrasco.
- c. V P F Alguns que foram ao churrasco não são alunos da professora Ana.
- d. V P F Alguns alunos da professora Ana não foram ao churrasco.

Para as duas situações abaixo, conclua conforme as três possibilidades:

- V necessariamente verdadeiro
- P não é possível ter certeza
- F necessariamente falso

8. Situação: Todo elemento pertencente ao conjunto **A** pertence ao conjunto **B**. Então: Você pode concluir que (assinale V, P ou F em cada uma das conclusões):

- a. V P F Todo elemento que pertence a **B** também pertence a **A**.
- b. V P F Algum elemento que pertence a **B** também pertence a **A**.
- c. V P F Algum elemento que pertence a **B** não pertence a **A**.
- d. V P F Nenhum elemento que pertence a **B** pertence a **A**.

9. Situação: Algum elemento pertencente ao conjunto **A** não pertence ao conjunto **B**. Então:

Você pode concluir que (assinale V, P ou F em cada uma das conclusões):

- a. V P F Todo elemento que pertence a **B** também pertence a **A**.
- b. V P F Algum elemento que pertence a **B** também pertence a **A**.
- c. V P F Algum elemento que pertence a **B** não pertence a **A**.
- d. V P F Nenhum elemento que pertence a **B** pertence a **A**.

As questões 10 a 16 referem-se ao texto abaixo:

Em um jogo há exatamente seis latas de boca para baixo, enumeradas consecutivamente de 1 a 6, colocadas lado a lado em linha reta. Embaixo de cada lata há uma bola escondida. Cada bola é de uma cor: verde, azul, laranja, lilás, vermelha e amarela. As bolas estão escondidas conforme as seguintes regras:

- A bola lilás está escondida embaixo de uma lata que tem um número inferior à lata que esconde a bola laranja.
- A bola vermelha está escondida embaixo de uma lata imediatamente ao lado da lata que esconde a bola azul.
- A bola verde está escondida embaixo da lata 5.

10. Qual alternativa abaixo descreve uma ordem correta para as cores das bolas debaixo das latas 1 a 6?

- a. verde, amarela, azul, vermelha, lilás, laranja
- b. azul, verde, lilás, vermelha, laranja, amarela
- c. azul, vermelha, lilás, amarela, verde, laranja
- d. laranja, amarela, vermelha, azul, verde, lilás
- e. vermelha, lilás, azul, amarela, verde, laranja

11. Se a bola de cor azul estiver embaixo da lata 4, a bola vermelha deve estar debaixo da lata:

- a. 1
- b. 2
- c. 3
- d. 5
- e. 6

12. De que cor poderia ser a bola embaixo da lata 6?

- a. verde
- b. azul
- c. lilás
- d. vermelha
- e. amarela

13. Se a bola lilás estiver embaixo da lata 4, a bola laranja deverá estar debaixo da lata:

- a. 1
- b. 2
- c. 3
- d. 5
- e. 6

14. Qual das seguintes afirmações não poderia ser verdadeira?

- a. A bola verde está embaixo de uma lata com um número menor do que a lata que esconde a bola amarela.
- b. A bola laranja está embaixo de uma lata com um número menor do que a lata que esconde a bola verde.
- c. A bola lilás está embaixo de uma lata com um número maior do que a lata que esconde a bola verde.
- d. A bola lilás está embaixo de uma lata com um número menor do que a lata que esconde a bola vermelha.
- e. A bola vermelha está embaixo de uma lata com um número menor do que a lata que esconde a bola amarela.

15. Se a bola laranja está embaixo da lata 2, de que cores poderiam ser as bolas embaixo das latas imediatamente adjacentes entre si?

- a. verde e azul
- b. verde e lilás
- c. laranja e amarela
- d. lilás e vermelha
- e. vermelha e amarela

16. Se a bola azul está debaixo da lata 1, de que cores poderiam ser as bolas embaixo das latas imediatamente adjacentes entre si (na seqüência)?

- a. verde e laranja
- b. verde e amarela
- c. lilás e vermelha

- d. lilás e amarela
- e. vermelha e amarela

As questões 17 a 25 referem-se ao texto abaixo:

Você é o juiz encarregado de julgar o caso que se apresenta abaixo, no qual o Cursinho Ômega, acusa o Cursinho Alfa de propaganda enganosa. Você recolheu os seguintes depoimentos:

- Cursinho Alfa: “Todos os nossos alunos passaram no vestibular.”
- Carla: “Conheço uma pessoa que estudou no Alfa e passou no vestibular.”
- Fernando: “Conheço todos os alunos do Alfa e sei que um dos alunos não passou no vestibular.”

- Cursinho Ômega: “Nenhum aluno do Alfa passou no vestibular.”

Apesar de não ser possível saber quem está falando a verdade, você pode concluir que:

Se Fernando disse a verdade, então:	17. Carla estaria:	a. Certa b. Errada c. Não é possível julgar
	18. “Alfa” estaria:	a. Certo b. Errado c. Não é possível julgar
	19. “Ômega” estaria:	a. Certo b. Errado c. Não é possível julgar
Se Carla disse a verdade, então:	20. “Alfa” estaria:	a. Certo b. Errado c. Não é possível julgar
	21. Fernando estaria:	a. Certo b. Errado c. Não é possível julgar
	22. “Ômega” estaria:	a. Certo b. Errado c. Não é possível julgar
Se “Ômega” disse a verdade, então:	23. Carla estaria:	a. Certo b. Errado c. Não é possível julgar
	24. Fernando estaria:	a. Certo b. Errado c. Não é possível julgar
	25. “Alfa” estaria:	a. Certo b. Errado c. Não é possível julgar

ANEXO B

Declaramos para os devidos fins que aceitaremos o pesquisador Mozart Amorim Sousa, a desenvolver o seu projeto de pesquisa do mestrado em psicologia cognitiva: **AValiação DO PENSAMENTO CRÍTICO: ANÁLISE DAS PROPRIEDADES PSICOMÉTRICAS DE UM TESTE DE RACIOCÍNIO LÓGICO-DEDUTIVO**, que está sob a coordenação/orientação do Prof. José Maurício Haas Bueno, cujo objetivo é: aplicação (e posterior estudo correlacional) em uma população nordestina dos testes de Raciocínio Lógico-Dedutivo (Primi & Muller, não publicado); aplicação da Prova de Raciocínio Abstrato (RA), Forma B, da Bateria de Provas de Raciocínio (BPR-5), que é composta por 25 itens e por fim a aplicação da Bateria Fatorial de Personalidade (BFP), que é baseada no Modelo dos Cinco Grandes Fatores.

A aceitação está condicionada ao cumprimento do pesquisador aos requisitos da Resolução 466/12 e suas complementares, comprometendo-se a utilizar os dados e materiais coletados, exclusivamente para os fins da pesquisa.

Recife, em ____/____/_____.

Nome/assinatura e carimbo do responsável pela Instituição

ANEXO C

UNIVERSIDADE FEDERAL DE PERNAMBUCO
TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO
(PARA MAIORES DE 18 ANOS OU EMANCIPADOS - RESOLUÇÃO 466/12)

Convidamos o (a) Sr. (a) para participar, como voluntário(a), da pesquisa **AVALIAÇÃO DO PENSAMENTO CRÍTICO: ANÁLISE DAS PROPRIEDADES PSICOMÉTRICAS DE UM TESTE DE RACIOCÍNIO LÓGICO-DEDUTIVO**, que está sob a responsabilidade do (a) pesquisador (a) Mozart Amorim Sousa, Rua Prof. Arthur de Sá, s/n, CEP: 50670-901, Cidade Universitária; Telefone 081 8802-6864 (inclusive ligações a cobrar); e-mail: mozart_fono@hotmail.com e está sob a orientação de: José Maurício Haas Bueno, Telefone 081 9772-9331 (inclusive ligações a cobrar); e-mail: talktomau@gmail.com.

Após ser esclarecido (a) sobre as informações a seguir, no caso de aceitar a fazer parte do estudo, o (a) Sr.(a) deve assinar este termo concordando com os procedimentos desta pesquisa. Em caso de recusa o (a) Sr.(a) não será penalizado (a) de forma alguma. O (a) Senhor (a) tem o direito de retirar o consentimento a qualquer tempo, sem qualquer penalidade.

INFORMAÇÕES SOBRE A PESQUISA

A presente pesquisa pretende coletar dados em formulário impresso através de perguntas com situações-problema, onde será solicitado que o participante analise as possibilidades de resposta escolhendo a que considerar mais apropriada na resolução da questão, com objetivo de avaliar formas de raciocínio utilizadas pelo mesmo e classificando as respostas em fatores que subjazem o pensamento crítico.

Esta pesquisa contará com a participação de voluntários somente durante o período de resposta aos testes.

Os riscos desta pesquisa estão relacionados ao possível constrangimento em virtude de não saber ou não querer responder ao questionário. Esses riscos poderão ser minimizados com a interrupção imediata da pesquisa pelo participante.

Como benefício direto o participante receberá, ao final completo da pesquisa, informações acerca do seu desempenho nos testes. Como benefício indireto o participante estará contribuindo para a produção de conhecimento científico que poderá auxiliar na melhoria dos processos educacionais. As informações desta pesquisa serão confidenciais e poderão ser divulgadas apenas em eventos ou publicações científicas, não havendo em nenhuma circunstância a identificação dos voluntários, a não ser entre os próprios responsáveis pelo estudo, sendo ainda assegurado o sigilo sobre a sua participação e seus dados pessoais. Os dados coletados nesta pesquisa (respostas ao questionário e aos testes), ficarão armazenados em pastas de arquivo e seus dados digitalizados serão posteriormente salvos em formato de CD, sob cuidados do pesquisador principal, o qual pode ser encontrado no endereço acima descrito pelo período mínimo 5 anos.

O (a) senhor(a) não pagará nada para participar desta pesquisa. Se houver necessidade, as despesas para a sua participação serão assumidas pelos pesquisadores (ressarcimento de despesas). Fica também garantida indenização em casos de danos, comprovadamente decorrentes da participação na pesquisa, conforme decisão judicial ou extra-judicial.

Em caso de dúvidas relacionadas aos aspectos éticos deste estudo, você poderá consultar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: (Avenida da Engenharia

s/n –1º Andar, sala 4 - Cidade Universitária, Recife-PE, CEP: 50740-600, Tel.: (81) 2126.8588 – e-mail: cepccs@ufpe.br).

CONSENTIMENTO DA PARTICIPAÇÃO DA PESSOA COMO VOLUNTÁRIO (A)

Nome completo (letra de forma):

Eu, _____

Portador(a) do RG _____ e do CPF _____ concordo em participar do estudo intitulado AVALIAÇÃO DO PENSAMENTO CRÍTICO: ANÁLISE DAS PROPRIEDADES PSICOMÉTRICAS DE UM TESTE DE RACIOCÍNIO LÓGICO-DEDUTIVO, como voluntário(a). Fui devidamente informado (a) e esclarecido (a) sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes de minha participação. Foi-me garantido que posso retirar meu consentimento a qualquer momento, sem que isto leve a qualquer penalidade.

Local: _____

Data: _____

Ass. _____

Presenciamos a solicitação de consentimento, esclarecimentos sobre a pesquisa e o aceite do voluntário em participar. (02 testemunhas não ligadas à equipe de pesquisadores):

Testemunha

Nome: _____

Assinatura: _____

Testemunha

Nome: _____

Assinatura: _____