

Teste do Pensamento Crítico e Criativo para estudantes do ensino superior

José Lopes, Helena Silva & Eva Morais

Resumo:

Não existindo um teste de avaliação do pensamento crítico pensado e construído de raiz para ser utilizado em Portugal, o foco deste artigo é o Teste do Pensamento Crítico e Criativo (TPCC), para jovens adultos portugueses. Apresentamos o estudo exploratório de validação deste teste, administrado a 250 estudantes universitários, com idades entre os 17 e os 37, a frequentar uma licenciatura ou mestrado STEM (Ciência, Tecnologia, Engenharia e Matemática) e de Ciências Humanas e Sociais (CHS). Da Análise Fatorial Exploratória emergiram três fatores (criação, argumentação e análise), que ilustram componentes essenciais do pensamento crítico. O TPCC revelou deter boas características psicométricas de validade e fiabilidade dos constructos (alpha de Cronbach igual a 0.68) e fiabilidade inter-juízes (kappa de Cohen entre 0.76 e 0.93 para cada dimensão). Verificou-se que os estudantes de mestrado obtiveram pontuações significativamente mais elevadas no TPCC do que os dos três anos das licenciaturas, e que as pontuações dos estudantes de STEM foram mais elevadas do que as pontuações dos de CHS. Não foram encontradas diferenças estatisticamente significativas nas pontuações médias no TPCC em função do género.

Palavras-chave:

pensamento crítico; ensino superior; avaliação; validação.

Critical and Creative Thinking Test for Higher Education Students

Abstract: As there is no critical and creative thinking test built from scratch to be used in Portugal, the focus of this work is the Critical and Creative Thinking Test for Portuguese young adults. We present the exploratory study of validation of this test, applied to 250 university students, aged from 17 to 37, attending a STEM (Science, Technology, Engineering and Mathematics) or a Human and Social Sciences degree (HSS). From the Exploratory Factor Analysis emerged three factors (creation, argumentation and analysis), which illustrate essential components of critical thinking. The Critical and Creative Thinking Test revealed good psychometric characteristics of constructs validity and reliability (Cronbach's alpha equal to 0.68) and interrater reliability (Cohen's kappa between 0.76 and 0.93 for each dimension). Students of a Master's degree were found to have scored significantly higher on the Critical and Creative Thinking Test than students of a Bachelor's degree, and the scores for STEM students were higher than the scores of HSS students. No statistically significant differences were found in the mean scores of the Critical and Creative Thinking Test according to gender.

Keywords: critical thinking; higher education; assessment; validation.

Test sur la pensée critique et créative pour les étudiants de l'enseignement supérieur

Résumé: Comme il n'existe aucun test de pensée critique et créative construit à partir de zéro pour être utilisé en Portugal, le présent travail est centré sur le Test de pensée critique et créative (TPCC) pour les jeunes adultes portugais. Nous présentons l'étude exploratoire de validation de ce test, appliquée à 250 étudiants universitaires, de 17 à 37 ans, fréquentant un cours STEM (Sciences, Technologie, Ingénierie et Mathématiques) ou de Sciences Humaines et Sociales (SHS). L'analyse exploratoire des facteurs a permis de dégager trois facteurs (création, argumentation et analyse) illustrant les éléments essentiels de la pensée critique. Le TPCC a révélé de bonnes caractéristiques psychométriques de validité et de fiabilité des construits (alpha de Cronbach égal à 0.68) et d'arbitrer la fiabilité (kappa de Cohen compris entre 0.76 et 0.93 pour chaque dimension). Les étudiants de maîtrise ont obtenu des résultats significativement plus élevés au TPCC que les étudiants de licence et les scores des étudiants en STEM sont supérieurs à ceux de la SHS. Aucune différence statistiquement significative n'a été trouvée dans les scores moyens du TPCC en fonction du sexe.

Mots-clés: pensée critique; enseignement supérieur; évaluation; validation.

Prueba del Pensamiento Crítico y Creativo para estudiantes de enseñanza superior

Resumen: Como no hay una prueba de pensamiento crítico y creativo creada desde cero para ser utilizada en Portugal, el enfoque de este trabajo es la Prueba de pensamiento crítico y creativo (TPCC) para adultos jóvenes portugueses. Presentamos el estudio exploratorio de validación de esta prueba, aplicado a 250 estudiantes universitarios, de entre 17 y 37 años, que cursan un grado STEM (Ciencia, Tecnología, Ingeniería y Matemáticas) o de Ciencias Humanas y Sociales (CHS). Del análisis factorial exploratorio surgieron tres factores (creación, argumentación y análisis), que ilustran los componentes esenciales del pensamiento crítico. El TPCC reveló buenas características psicométricas de la validez y confiabilidad de los constructos (alfa de Cronbach igual a 0.68) y confiabilidad entre los jueces (kappa de Cohen entre 0.76 y 0.93 para cada dimensión). Se encontró que los estudiantes de una maestría obtuvieron puntuaciones significativamente más altas en el TPCC que los estudiantes de una licenciatura, y las puntuaciones para los estudiantes STEM fueron más altas que las puntuaciones de los estudiantes CHS. No se encontraron diferencias estadísticamente significativas en las puntuaciones medias en el TPCC según el género.

Palabras clave: pensamiento crítico; enseñanza superior; evaluación; validez.

Introdução

A importância do pensamento crítico (PC) é bem fundamentada na literatura. Apesar da heterogeneidade de definições para descrever este constructo, bem como de modelos teóricos para o explicar, encontramos consenso no que concerne à sua relevância, não só académica, como profissional, individual e, ainda, social. As competências e disposições de PC estão associadas ao sucesso académico (Halpern, 2014). Também no mercado de trabalho, o PC encontra-se no topo das competências mais pretendidas pelas empresas (Brotherton, 2011; Weng, 2015). Adicionalmente, para além de potenciar o sucesso académico e profissional, o PC apresenta-se como uma vantagem na vida pessoal e social do dia a dia. Segundo os estudos de Butler et al. (2012), um nível superior de PC está associado à tomada de decisões mais informadas e fundamentadas, ou seja, decisões reflexivas que ajudam a evitar um leque de ocorrências negativas resultantes da imprudência ou impulsividade. As competências e disposições que constituem o núcleo-duro do PC nutrem uma postura ativa, participativa e ética da parte do indivíduo, pelo que o PC se encontra associado, na literatura, a uma mais expressiva ação da cidadania (OCDE, 2015).

Dada a sua relevância, são inúmeros os trabalhos de teorização, avaliação e de intervenção na área do PC, nacional e internacionalmente, visando a análise, medição, e/ou promoção do PC nos diversos níveis de escolaridade e áreas de formação académica. No que respeita à sua avaliação, em particular, não existe até ao momento um teste de avaliação do pensamento crítico pensado e construído de raiz para ser utilizado em Portugal e por portugueses. Daqui resultam restrições para quem, no contexto da investigação e/ou da prática, deseje medir o PC para melhor o compreender, quantificar e/ou promover, dado que é assumido que as competências de PC podem ser ensinadas e aprendidas pelos alunos dos vários níveis de ensino. Por estes motivos, é importante que professores e alunos tenham uma consciência clara do nível de domínio das mesmas no decurso do processo de ensino e aprendizagem.

Pensamento crítico: definições

Existe uma variedade de definições para caracterizar o PC (Sternberg, Roediger, & Halpern, 2007). No entanto, mais do que se dever a uma eventual incongruência entre autores e suas definições, tal riqueza conceptual parece derivar das múltiplas faces deste constructo, com cada autor optando por se focar numa faceta do PC em particular. De acordo com Norris e Ennis (1989, p.1) é um “pensamento racional e reflexivo, centrado em definir em que acreditar ou no que fazer”. Para Halpern (1998, p. 450) é a utilização de competências cognitivas ou estratégias que aumentam a probabilidade de obter resultados desejados”. Facione (1990b, p.21) define-o como “o julgamento

intencional, autorregulado que resulta em interpretação, análise, avaliação e inferência assim como explicação das considerações de evidências e conceituais, metodológicas, criteriológicas ou contextuais, sobre as quais o julgamento foi baseado”. Em traços gerais, pensamento crítico pode ser definido como “uma forma superior de pensamento integrando competências cognitivas e disposições, ‘feito à medida’, pois aplica-se no dia a dia consoante a situação para aumentar a possibilidade de se alcançar o objetivo pretendido” (Franco, 2016, p. 21).

Uma vez que o PC tem uma natureza funcional, agregando em si competências relativas ao funcionamento cognitivo ligado à resolução de problemas, não raras vezes se fala de criatividade, ou pensamento criativo, ao falar-se de pensamento crítico. Para Paul e Elder (2006) a criatividade e o pensamento crítico são como os dois lados da mesma moeda. Ou seja, na prática, os dois conceitos estão intrinsecamente ligados desenvolvendo-se em paralelo. O bom pensamento requer a capacidade de gerar produtos intelectuais, o que está associado à criatividade. Como observam os autores, “o pensamento crítico sem criatividade reduz-se ao mero ceticismo e negatividade, e a criatividade sem pensamento crítico reduz-se a mera novidade” (p. 35). A criatividade encerra em si características como a flexibilidade, a apetência para a resolução de problemas, ou para o trabalho e aprendizagem colaborativos (Hamalainen & Vahasantanen, 2011).

A avaliação do pensamento crítico

A nível internacional existe um conjunto de testes de referência para medir PC, em função do grau de escolaridade, idade e propósito específico de avaliação. São exemplos, Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 2009), Ennis-Weir Critical Thinking Essay Test (Ennis & Weir, 1985), The California Critical Thinking Skills Test: College Level (Facione, 1990a), Cornell Critical Thinking Test - Level X (Ennis & Millman, 1985), Cornell Critical Thinking Test - Level Z (Ennis & Millman, 1985), Halpern Critical Thinking Assessment, (Halpern, 2012). Não há um teste de avaliação pensado e construído de raiz para ser utilizado em Portugal que meça o PC e que possa ser administrado a estudantes no ensino superior.

Para superar tal lacuna, têm sido levados a cabo estudos de tradução de testes utilizados internacionalmente, como é o caso do *Cornell Critical Thinking Test - Level X* (Ennis & Millman, 1985), designado de *Teste de Pensamento Crítico de Cornell - Nível X* (Oliveira, 1992) na sua versão traduzida. Ainda adaptado para português (Rocha, Ferreira, & Oliveira, 2004) encontra-se disponível o *GMA - Avaliação de Quadros Médios e Superiores* (Forma B) (Blinkhorn, 2003). Mais recentemente, traduziu-se, adaptou-se e validou-se o *Halpern Critical Thinking Assessment* (Halpern, 2012) para estudantes universitários portugueses (Franco, Costa, & Almeida, 2018), tendo restrições editoriais.

Aliado a este facto, há uma falta evidente de estudos experimentais e longitudinais ao nível do impacto de estratégias de ensino no desenvolvimento do PC a nível nacional, que pode derivar sobretudo da existência de barreiras e limitações no uso dos testes existentes por parte dos docentes (Cruz, Payan-Carreira, & Dominguez, 2017).

Perante esta realidade e dadas as restrições daqui resultantes para quem, no contexto da investigação ou da prática, deseja medir as competências de pensamento crítico dos estudantes universitários portugueses para melhor as compreender, quantificar e/ou promover, foi objetivo deste estudo construir e validar uma escala de PC adaptada à população portuguesa, colmantando, assim, a lacuna existente.

Apresentamos, em seguida, o estudo exploratório de validação do Teste do Pensamento Crítico e Criativo (TPCC).

O pensamento crítico em estudo: Teste do Pensamento Crítico e Criativo

Construído originalmente para alunos no ensino superior, o TPCC representa um empreendimento para se obter um teste em língua portuguesa que possa ser utilizado para avaliar, de forma compreensiva, o pensamento destes estudantes. Este teste propõe uma situação-problema da vida real para avaliar competências do PC com base na taxonomia revista de Bloom (Anderson & Krathwohl, 2001) e na taxonomia de Facione (1990b).

A Taxonomia revista de Bloom identifica três grandes domínios de desenvolvimento – cognitivo, afetivo e psicomotor –, todos eles com níveis cumulativos gradualmente mais complexos. Neste artigo, iremos debruçar-nos sobre o domínio cognitivo de desenvolvimento que apresenta as seguintes categorias: (1) *lembrar*, (2) *entender*, (3) *aplicar*, (4) *analisar*, (5) *sintetizar*, (6) *criar*. Analisar, sintetizar e criar são competências de pensamento crítico (Kennedy, Fisher, & Ennis, 1991). Esta taxonomia é relevante para os professores no planeamento educacional para que os objetivos de aprendizagem visados, as estratégias a serem utilizadas e os formatos de avaliação que serão adotados sejam claros para os alunos, o que facilita a sua aprendizagem.

No que diz respeito à taxonomia proposta por Facione (1990b), esta surge no contexto do consenso alcançado utilizando o método de Delphi, com o objetivo de determinar quais os constituintes nucleares do PC, assim como as abordagens mais eficazes para se ensinar e se avaliar este constructo. Neste relatório são descritas seis dimensões centrais do PC: interpretação, análise, avaliação, inferência, explicação e autorregulação. A *interpretação* diz respeito aos processos de categorização, de descodificação do significado e de clarificação de sentido. A *análise* diz respeito à análise de ideias, à identificação de argumentos, bem como à análise dos mesmos. A *avaliação* relaciona-se com a avaliação de suposições e de argumentos. A *inferência* respeita a consulta de provas, a ponderação de alternativas e a elaboração de

conclusões. A *explicação* tem que ver com a apresentação de resultados, a explanação de procedimentos e a apresentação de argumentos. A *autorregulação* descreve a autoavaliação, assim como a autocorreção. Na implementação de cada competência, estão implicadas disposições ou atitudes que fazem o pensador crítico querer fazer uso das suas competências, as quais o parecem perfilar perante a vida e o viver em geral, bem como face a problemas e/ou situações específicas.

Facione (2011) considera que o PC está intrinsecamente ligado à criatividade, mais especificamente à componente criativa. Para criar, os alunos têm de elaborar um novo produto ou soluções diferentes dos materiais ou situações originais. Para Guilford (1986), a criatividade integra quatro componentes essenciais: fluência, flexibilidade, originalidade e elaboração.

Com base nas três taxonomias apresentadas iniciou-se o projeto de construção e validação de um teste para a avaliação do PC que pudesse também ser utilizado em situações de pré e pós-teste, para a promoção do PC em sala de aula para alunos universitários. O teste tem um formato de resposta aberta sobre um problema da vida real, dado que as questões de resposta aberta permitem que os alunos vão mais além da informação disponível na situação problema proposta para fazerem inferências ou avaliações (Ku, 2009; Moss & Koziol, 1991). Além disso, os problemas da vida real são problemas abertos que, tendo mais do que uma solução plausível ou defensável, exigem que os alunos fundamentam as suas tomadas de decisão (Bonk & Smith, 1998).

São objetivos deste estudo a construção e validação de um teste de avaliação do pensamento crítico e criativo (TPCC) em língua portuguesa que permita avaliar as competências de pensamento crítico e criativo de estudantes universitários e investigar se o ciclo de estudos que frequentam têm influência nas competências de PC.

Método

Participantes

Participaram 250 estudantes de uma universidade pública portuguesa, com idades compreendidas entre os 17 e os 37 anos ($M=21.1$, $DP=3.0$). As áreas curriculares frequentadas eram STEM, na qual se incluem cursos de Engenharia e Gestão Industrial, Ciência Alimentar, Medicina Veterinária, Engenharia Informática, Engenharia Mecânica e Bioengenharia (40.8% dos participantes), assim como CHS, na qual se incluem cursos de Psicologia, Educação, Línguas e Relações Internacionais, e Reabilitação Psicomotora (59.2% da amostra). Da amostra, 91 (36.4%) participantes eram do género masculino e 159 (63.6%) eram do género feminino. Destes, 70.3% dos participantes do género masculino encontravam-se a frequentar um curso na área de STEM e 76.1% dos participantes do género feminino frequentavam um curso na área de CHS. A distribuição por ciclo de graduação foi de 34% de alunos no 1.º ano de licenciatura, 18.4%

no 2.º ano de licenciatura, 23.6% no 3.º ano de licenciatura e, finalmente, 24% no 1.º ano de mestrado.

Instrumento

O TPCC apresenta, inicialmente, uma situação-problema atual e comum do quotidiano, na qual se pode identificar um conjunto de circunstâncias problemáticas. Ao respondente é pedido que responda a um conjunto de seis perguntas, as quais se referem a diferentes tarefas cognitivas associadas ao domínio do PC e que requerem o emprego das seguintes competências do PC para a sua apreciação e resolução: (i) interpretação, (ii) análise, (iii) explicação, (iv) avaliação, (v) síntese e (vi) produção/criação. Na pergunta número 1, é pedido ao respondente que identifique o problema contido na situação descrita, de modo a avaliar-se a sua capacidade de (i) *interpretação* – i.e., de compreender e expressar o significado de uma ampla variedade de experiências, situações, dados, acontecimentos, juízos, convenções, crenças, regras, procedimentos ou critérios. Na pergunta número 2, é pedido que o respondente identifique e, depois, compare soluções para a resolução do problema, para avaliar-se a sua capacidade de (ii) *análise* – i.e., de apresentar as ideias principais e relacioná-las entre si, o que envolve decompor o material nas suas partes constituintes e determinar como as partes se relacionam entre si e com a estrutura geral da informação. Na pergunta número 3, é pedido ao respondente que selecione a melhor solução e apresente um argumento em sua defesa, com vista a avaliar-se a sua capacidade de (iii) *explicação* – i.e., de apresentar o resultado do seu próprio raciocínio e justificar esse raciocínio com argumentos válidos. Na pergunta número 4, é pedido ao respondente que aprecie a qualidade da solução defendida, de modo a determinar-se a sua capacidade de (iv) *avaliação* – i.e., para avaliar a credibilidade de argumentos, representações, descrições das perceções, experiências, situações, avaliações, crenças ou opiniões pessoais, assim como das relações inferenciais entre argumentos, descrições, questões ou outras formas de representação. Na pergunta número 5, é pedido ao respondente que proponha estratégias para maximizar a qualidade da solução por si defendida, para avaliar-se a sua capacidade de (v) *síntese* – i.e., de construir conhecimento com base na recolha e tratamento da informação, sendo necessário utilizar informação para melhorar a solução apresentada ou defendida. Na pergunta número 6, é pedido ao respondente que apresente novas soluções, de forma imaginativa e inovadora, para a resolução do problema em mãos, para avaliar-se a sua capacidade de (vi) *produção/criação* – i.e., de apresentar uma ou mais soluções novas para o problema. Cada uma das seis perguntas colocadas deve ser respondida abertamente pelos respondentes numa folha de respostas.

Para cada uma das seis perguntas, o avaliador deverá pontuar as respostas de acordo com uma escala de 0 a 3 pontos, com a exceção das questões 2b e 6b que

são cotadas numa escala de 0 a 2 pontos. Os pontos atribuem-se em função da dimensão do pensamento crítico avaliada por cada pergunta, conforme a sua qualidade e quantidade (Tabela 1).

Tabela 1 - Grelha de atribuição de pontos a cada dimensão do PC avaliada

| Dimensão | Critérios de avaliação | Pontos |
|---------------------|--|--------|
| 1. Interpretação | Resposta completa, fundamentada e organizada. | 3 |
| | Resposta incompleta, pouco fundamentada e pouco organizada. | 2 |
| | Resposta incompleta, infundamentada e confusa. | 1 |
| | Resposta inteiramente incorreta. | 0 |
| 2. Análise | Identifica todas as soluções possíveis, ou então, pelo menos duas diferenças e uma semelhança entre as soluções. | 3 |
| | Identifica duas soluções possíveis, ou então, duas diferenças, ou ainda uma diferença e uma semelhança, entre as soluções. | 2 |
| | Identifica apenas uma solução possível, ou então, uma diferença, ou ainda, uma semelhança entre as soluções. | 1 |
| | Não identifica nenhuma solução, ou então, não estabelece nenhuma comparação entre as soluções. | 0 |
| 3. Explicação | O argumento é coerente e possui duas ou mais premissas justificativas. | 3 |
| | O argumento é coerente e possui uma premissa justificativa. | 2 |
| | O argumento é pouco coerente. | 1 |
| | O argumento é inválido. | 0 |
| 4. Avaliação | Resposta coerente e apresenta pelo menos três pontos fracos. | 3 |
| | Resposta coerente e/ou apresenta dois pontos fracos. | 2 |
| | Resposta pouco coerente e/ou apresenta pelo menos um ponto fraco. | 1 |
| | Resposta apresenta pontos inválidos ou não os apresenta. | 0 |
| 5. Síntese | O argumento melhora, de forma coerente, todos os pontos fracos de uma solução ou origina uma nova solução. | 3 |
| | O argumento melhora, de forma coerente, alguns dos pontos fracos de uma solução. | 2 |
| | O argumento é pouco coerente com os pontos fracos da solução em causa. | 1 |
| | O argumento não melhora os pontos fracos identificados. | 0 |

| Dimensão | Critérios de avaliação | | Pontos |
|----------------------------|------------------------|---|--------|
| 6. Produção/ Criação | Fluência | Apresenta mais de duas soluções. | 3 |
| | | Apresenta duas soluções. | 2 |
| | | Apresenta uma solução. | 1 |
| | | Não responde. | 0 |
| | Flexibilidade | As soluções apresentadas servem para a resolução do problema. | 2 |
| | | A solução apresentada serve para a resolução do problema, ou algumas das soluções servem para a resolução do mesmo. | 1 |
| | | Nenhuma das soluções serve para a resolução do problema, ou são iguais às propostas no texto. | 0 |
| | Originalidade | Pelo menos uma das soluções apresentadas é nova e as outras não têm por base pressupostos semelhantes às apresentadas, ou ainda, são menos referidas. | 3 |
| | | Das soluções apresentadas, as que servem para a resolução do problema são novas e muito referidas. | 2 |
| | | A solução (ou soluções) apresentada é uma modificação ou melhoria das propostas no texto. | 1 |
| | | A solução (ou soluções) apresentada é uma cópia das já propostas. | 0 |

A pergunta número 6, que avalia a dimensão de *produção/criação* do PC, foi avaliada mediante três das quatro componentes da criatividade enunciadas: fluência, flexibilidade, originalidade e elaboração. No TPCCC, o parâmetro da fluência avaliou o número de soluções apresentadas para o problema; o da flexibilidade avaliou a variedade das soluções; a originalidade avaliou a novidade das ideias apresentadas para a resolução do problema (Tabela 1), fixada nos critérios estabelecidos para a atribuição de pontos. Não foi considerada a componente elaboração porque a pergunta não induzia a que os respondentes detalhassem as soluções.

Procedimentos

Recolha de dados

Foram contactados professores universitários para obter-se autorização de aplicação do instrumento nas suas turmas. Os estudantes responderam ao teste em contexto de sala de aula, estando o docente presente no momento da aplicação. Os objetivos do estudo foram apresentados, tendo-se assegurado a participação voluntária dos estudantes, assim como o seu anonimato. O TPCCC demorou cerca de 30 minutos a

ser respondido por cada estudante. Quanto à cotação do teste, esta despendeu entre 5 a 10 minutos e foi levada a cabo por quatro juizes/avaliadores devidamente treinados nesta tarefa.

Análises estatísticas

Seguida da apreciação da normalidade, procedeu-se ao estudo de fiabilidade através da análise da consistência interna do teste e do acordo inter-juizes, tendo-se estimado os coeficientes kappa de Cohen para os 250 testes. Estimou-se a validade de constructo mediante uma Análise Fatorial Exploratória (AFE), para avaliar a multidimensionalidade do TPCC e examinar se a expectativa teórica da existência de três fatores no instrumento se concretizava. Tratando-se da criação de um novo instrumento, interessa compreender o número e a natureza dos fatores que contribuem para a explicação da variância e covariância entre um conjunto de indicadores observados ou itens (Brown, 2006). Realizaram-se testes *t* para explorar a hipótese de diferenças no teste de acordo com: género, ano de escolaridade (1.º, 2.º e 3.º anos de licenciatura e 1.º de mestrado), ciclo de graduação (licenciatura e mestrado) e área curricular (STEM e CHS). As análises quantitativas foram realizadas recorrendo ao programa IBM SPSS, versão 22.0 para *Windows*.

Resultados

As pontuações totais dos 250 participantes no TPCC apresentam uma média de 13.12 (IC 95%: 12.61-13.63), com o desvio-padrão de 4.09 para um intervalo de pontuações: 2-22 (a pontuação máxima é 25 pontos). Uma vez assegurada a normalidade da distribuição das pontuações (por género, ano de escolaridade, ciclo de graduação e área disciplinar), procedeu-se ao estudo de fiabilidade do TPCC mediante a análise da consistência interna do teste e do acordo inter-juizes. A consistência interna dos seis itens que compõem o teste foi calculada pelo alpha de Cronbach; uma vez que o item 6 é classificado em três dimensões (fluência, flexibilidade e originalidade), avaliou-se um total de nove componentes do pensamento crítico. O coeficiente de fiabilidade obtido foi de .68, considerado altamente significativo ($n=250$, $F(8;1991)=55.02$, $p<.001$), revelando um grau de homogeneidade entre os itens aceitável. Para a fiabilidade inter-juizes, foram estimados os coeficientes kappa de Cohen cruzando-se os resultados dos 250 testes classificados por quatro juizes/avaliadores treinados nesta tarefa. Em todas as dimensões foram encontrados coeficientes superiores a .70 (Tabela 2), com valores considerados “bom” a “muito bom”, segundo o critério de Landis e Koch (1977).

Tabela 2 - Coeficiente de concordância inter-juizes

| Itens | Tarefa | Coeficiente kappa |
|-------|---|-------------------|
| 1 | Identifica o problema | .86 |
| 2a | Identifica as soluções | .91 |
| 2b | Compara as soluções | .86 |
| 3 | Soluciona e defende a solução escolhida | .76 |
| 4 | Identifica os pontos fracos da solução | .82 |
| 5 | Melhora a solução escolhida | .83 |
| 6a | Propõe novas soluções: Fluência | .93 |
| 6b | Propõe novas soluções: Flexibilidade | .89 |
| 6c | Propõe novas soluções: Originalidade | .83 |

O estudo da validade de constructo foi realizada com recurso à AFE (Costello & Osborne, 2005), utilizando o método de extração da Análise de Componentes Principais com rotação *varimax*. Para verificar se haveria a possibilidade de aplicar o método da AFE, realizou-se o Teste de Esfericidade de Bartlett, que indicou uma correlação entre os itens ($\chi^2=553.233$, $gl=36$, $p<.001$) portanto, a adequabilidade do uso da AFE. A medida de adequação da amostra foi averiguada pelo índice de Kaiser-Meyer-Olkin (KMO), de .725. A AFE indicou uma estrutura de três fatores para o teste, com *eigenvalues* acima de 1.0 capazes de explicar 61.98% da variância total (Tabela 3). O agrupamento dos fatores ficou assim distribuído: Fator 1 – *Criação*, com um item e três dimensões (item 6: Fluência, Flexibilidade e Originalidade); Fator 2 – *Argumentação*, com três itens (3, 4 e 5); Fator 3 – *Análise*, com três itens (1, 2a e 2b).

Tabela 3 - Fatores resultantes da AFE (Análise de Componentes Principais com rotação *varimax*)

| Itens | F1 | F2 | F3 | h2 |
|----------------------------|-------|-------|-------|------|
| Criatividade_Originalidade | .907 | | | .836 |
| Criatividade_Flexibilidade | .890 | | | .804 |
| Criatividade_Fluência | .862 | | | .755 |
| Avaliação | | .771 | | .603 |
| Síntese | | .712 | | .515 |
| Explicação | | .682 | | .539 |
| Análise_a | | | .729 | .547 |
| Análise_b | | | .676 | .561 |
| Interpretação | | | .639 | .419 |
| Eigenvalues | 2.74 | 1.69 | 1.15 | |
| % Variância explicada | 30.38 | 18.78 | 12.82 | |

Realizaram-se testes *t* para explorar a hipótese de diferenças no teste de acordo com as variáveis em estudo: género, ano de escolaridade, ciclo de graduação e área curricular. Adicionalmente, examinaram-se os resultados por item do teste em função do ciclo de graduação. Não foram encontradas diferenças estatisticamente significativas nas médias do teste em função do género ($t(248)=.247$; $p=.805$). Pelo contrário, foram encontradas diferenças estatisticamente significativas em função do ano de escolaridade, ($F(3,246)=13270$; $p<.001$), sendo que os estudantes do 1.º ano de licenciatura pontuaram menos no teste ($M=12.65$, $DP=4.10$) do que os do 1.º ano de mestrado ($M=15.57$, $DP=3.42$), tendo o mesmo acontecido no caso dos estudantes de 2.º ($M=11.02$, $DP=4.37$) e 3.º anos de licenciatura ($M=12.97$, $DP=3.28$) quando comparados com os de mestrado. Identificaram-se diferenças com significância estatística entre os dois ciclos de graduação ($t(248)=5.61$; $p<.001$), com os estudantes de mestrado a revelarem uma pontuação superior no teste ($M=15.57$, $DP=3.42$) do que os de licenciatura ($M=12.35$, $DP=3.99$). Ainda, foram detetadas diferenças estatisticamente significativas de acordo com a área curricular ($t(248)=2.68$; $p=.008$), com estudantes de STEM a pontuarem mais alto no teste ($M=13.95$, $DP=4.25$) do que os de CHS ($M=12.55$, $DP=3.89$). Por fim, examinaram-se os resultados em cada item do teste em função do ciclo de graduação (Tabela 4), com os estudantes de mestrado a pontuarem mais em todos os itens.

Tabela 4 - Resultados por item em função do ciclo de graduação dos participantes

| Itens | 1.º ciclo N=190 | | 2.º ciclo N=60 | |
|----------------------------|--------------------|------|-------------------|------|
| | M | DP | M | DP |
| Interpretação | 1.11 | .99 | 1.68 | 1.08 |
| Análise_a | 2.09 | .42 | 2.28 | .49 |
| Análise_b | .89 | .63 | 1.17 | .78 |
| Explicação | 1.65 | .93 | 2.03 | 1.04 |
| Avaliação | 1.15 | .76 | 1.63 | .92 |
| Síntese | 1.41 | 1.21 | 1.93 | 1.11 |
| Criatividade-Fluência | 1.11 | .70 | 1.40 | .78 |
| Criatividade-Flexibilidade | 1.24 | .72 | 1.42 | .74 |
| Criatividade-Originalidade | 1.69 | .93 | 2.05 | .91 |

Discussão e conclusões

O presente estudo visava a construção e validação de um teste de avaliação do pensamento crítico em língua portuguesa, a ser administrado a jovens estudantes, dada a sua relevância e as lacunas existentes no contexto de avaliação e intervenção do PC em Portugal. Esta versão experimental do TPCC foi administrada a uma amostra de estudantes universitários, primeiro para analisar a fiabilidade e validade do teste e depois para examinar a hipótese de diferenças entre os participantes de acordo com um conjunto de variáveis em estudo. As características psicométricas do teste, captadas mediante a análise da consistência interna e do acordo inter-juízes, revelam que é uma medida capaz de medir o constructo em causa. No que concerne a sua validade, da AFE emergiram três fatores – *criação*, *argumentação* e *análise* –, que ilustram três componentes essenciais do pensamento crítico. Por um lado, a *criação* relaciona-se com a dimensão de PC ligada à criatividade, que concerne características como a flexibilidade cognitiva, resolução de problemas, ou abertura à ideia de mudar de opinião (Franco & Almeida, 2017; Hamalainen & Vahasantane, 2011). Por outro lado, a *argumentação* e a *análise* são dimensões basilares do pensamento crítico. Descrevem a apetência para providenciar razões válidas e fundamentadas a favor de uma opção e ainda, de pensar reflexivamente sobre a informação, analisando a credibilidade da fonte e a qualidade das premissas de um argumento, assim como o que está expresso nas entrelinhas, de modo a deliberar sobre a validade das conclusões veiculadas pelo mesmo (Halpern, 2014). Deste modo, consideramos que o TPCC mede dimensões essenciais do pensamento crítico, como se pretendia.

A partir da análise de diferenças em função do género, ano de escolaridade, ciclo de graduação, e área curricular, observou-se que apenas o género não diferencia os participantes. Observando-se os resultados em cada um dos seis itens do TPCC, os estudantes de mestrado apresentaram sempre pontuações superiores aos estudantes de licenciatura. A relação entre pensamento crítico e ano de escolaridade tem sido o alvo de largos estudos, apontando-se uma relação proporcional entre o pensamento crítico e os anos de escolaridade, com pontuações superiores de pensamento crítico a surgirem em estudantes com mais anos de formação académica (Butler et al., 2012; Franco et al., 2018). Tal poderá dever-se ao impacto cumulativo das aprendizagens e à maior diversidade de experiências vividas, afetando, também, a qualidade do pensamento crítico, produto esperado do ensino, particularmente do ensino superior (Halpern, 2014).

No que respeita à área curricular, os estudantes a frequentar um curso STEM pontuaram mais no TPCC do que os de CHS, à semelhança de outros estudos (e.g., Franco et al., 2018). A questão da existência de diferenças no pensamento crítico entre estudantes de áreas distintas não é recente na literatura, havendo resultados apontando

numa e noutra direção (Franco, 2016). Contudo, as ditas diferenças de área curricular poderão derivar, não da área curricular em si, mas de características individuais dos próprios estudantes, que os influenciam a percorrer determinado percurso académico, mais voltado para as STEM ou para as CHS. Ainda, essas diferenças podem advir, não das características individuais dos estudantes, mas dos contextos de aprendizagem e de conteúdo associados a cada área (Franco et al., 2018). Face às hipóteses, é necessário dar continuidade aos estudos e examinar este tipo de variável, de modo a melhor determinar o que estará subjacente a este tipo de resultados.

Consideramos que este estudo, ainda que preliminar, é um marco no panorama nacional português no que concerne à avaliação do pensamento crítico e criativo e que futuros estudos são necessários para se criarem oportunidades de mensuração precisa e contextualizada destes constructos. Para além de não haver testes desta natureza, os que existem apresentam limitações como direitos editoriais, falta de contextualização e aplicação morosa. O TPCC é de acesso aberto e permite avaliar o pensamento crítico e criativo de forma holística. Além disso, tem as vantagens de poder ser aplicado no decurso de uma aula e corrigido de forma fácil e rápida por docentes, podendo os resultados ser usados para planificar atividades da promoção do pensamento crítico e criativo na sala de aula.

Em trabalhos futuros, consideramos necessário investigar como as competências e as disposições do pensamento crítico estão relacionadas, bem como se existe uma correlação com o desempenho escolar.

Dado que foi um estudo exploratório, e para uma maior consistência na validação e extensão à comunidade portuguesa, será importante incluir estudantes de outras universidades e áreas do conhecimento e países como os PALOP ou da CPLP.

Agradecimentos

Este trabalho insere-se no Projeto 'Pensamento Crítico nos Currículos do Ensino Superior Europeu – CRITINKEDU', com o número de referência 2016-1-PT01-KA203-022808, financiado pela Comissão Europeia no âmbito do Programa Erasmus+.

Referências Bibliográficas

- Anderson, L., & Krathwohl, D. (Eds.) (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York, NY: Longman.
- Blinkhorn, S. (2003). *Graduate and Managerial Assessment*. St. Albans, UK: Psychometric Research & Development Ltd.
- Bonk, C., & Smith, G. (1998). Alternative instructional strategies for creative and critical thinking in the accounting curriculum. *Journal of Accounting Education*, 16(2), 261–293. [https://doi.org/10.1016/S0748-5751\(98\)00012-8](https://doi.org/10.1016/S0748-5751(98)00012-8).

- Brotherton, P. (2011). Critical thinking: A top skill for future leaders. *T+D*, 65(11), 24.
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Butler, H., Dwyer, C., Hogan, M., Franco, A., Rivas, S., Saiz, C., & Almeida, L. (2012). The Halpern Critical Thinking Assessment and Real-World Outcomes: Cross-national applications. *Thinking Skills & Creativity*, 7(2), 112-121.
- Costello, A., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, 10(7), 1-9.
- Cruz, G., Payan-Carreira, R., & Dominguez, C. (2017). Critical thinking education in the Portuguese higher education institutions: a systematic review of educational practices. *Revista Lusófona de Educação*, 38, 43-61. doi: 10.24140/issn.1645-7250.rle38.03 [Disponível em <http://revistas.ulusofona.pt/index.php/rleducacao/article/view/6259>, consultado em 02-09-2018].
- Ennis, R., & Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test*. Pacific Grove, CA: Midwest Publications.
- Facione, P. (1990a). *The California Critical Thinking Skills Test: College Level*. Millbrae, CA: California Academic Press.
- Facione, P. (1990b). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report)*. Millbrae, CA: The California Academic Press.
- Facione, P. (2011). *Think critically*. Englewood Cliffs, NJ: Pearson Education.
- Franco, A. (2016). *Pensamento crítico: Desenvolvimento e mediação no quadro das experiências académicas no Ensino Superior*. Tese de doutoramento. Braga: Universidade do Minho.
- Franco, A., & Almeida, L. (2017). Definição e medida do pensamento crítico. In L. S. Almeida (Coord.), *Criatividade e pensamento crítico: Conceito, avaliação e desenvolvimento* (pp. 107-132). Porto: CERPSI.
- Franco, A., Costa, P., & Almeida, L. (2018). Translation, adaptation, and validation of the Halpern Critical Thinking Assessment to Portugal: Effect of disciplinary area and academic level on critical thinking. *Anales de Psicología*, 34(2), 292-298.
- Halpern, D. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449-455.
- Halpern, D. (2012). *Halpern Critical Thinking Assessment*. Mödling, Austria: Schuhfried.
- Halpern, D. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). New York, NY: Psychology Press.
- Hamalainen, R., & Vahasantani, K. (2011). Theoretical and pedagogical perspectives on orchestrating creativity and collaborative learning. *Educational Research Review*, 6(3), 169-184.
- Kennedy, M., Fisher, M., & Ennis, R. (1991). Critical thinking: Literature review and needed research. In L. Idol & B. Jones (Eds.), *Educational values and cognitive instruction: Implications for reform* (pp. 11-40). Hillsdale, New Jersey: Lawrence Erlbaum & Associates.
- Ku, K. Y. (2009). Assessing students' critical thinking performance: urging for measurements using multi-response format. *Thinking Skills and Creativity*, 4, 70-76.

- Landis, J., & Koch, G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 33(1), 363-374.
- Moss, P. Koziol, S. (1991). Investigating the validity of a locally developed critical thinking test. *Educational Measurement: Issues and Practice*, 10(3), 17-22. <https://doi.org/10.1111/j.1745-3992.1991.tb00199.x>
- Norris, S., & Ennis, R. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Midwest Publication.
- Oliveira, M. (1992). *A criatividade, o pensamento crítico e o aproveitamento escolar dos alunos de Ciências*. Tese de doutoramento. Lisboa: Universidade de Lisboa.
- Paul, R., & Elder, L. (2006). Critical thinking: The nature of critical and creative thought. *Journal of Developmental Education*, 30(2), 34-35. [Disponível em <https://ncde.appstate.edu/sites/ncde.appstate.edu/files/JDE%20TOC%20%20Abstract%20Website.pdf>, consultado em 20-12-2018].
- Rocha, A., Ferreira, C., & Oliveira, S. (2004). *GMA - Avaliação de quadros médios e superiores (Forma B)*. Lisboa, Portugal: Cegoc-TEA.
- Sternberg, R., Roediger, H. L., III, & Halpern, D. (2007). *Critical thinking in psychology*. New York, NY: Cambridge University Press.
- Watson, G., & Glaser, E. (2009). *Watson-Glaser Critical Thinking Appraisal*. London: Pearson Assessment.
- Weng, W. (2015). Eight skills in future work. *Education*, 135(4), 419-422. [Disponível em <https://www.questia.com/library/journal/1G1-418342192/eight-skills-in-future-work>, consultado em 25/6/2018].

José Lopes

Professor Associado

Universidade de Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal

Centro de Investigação e Intervenção Educativas (CIIE), Universidade do

Porto, Portugal

Email: jlopes@utad.pt

ORCID: 0000-0002-6845-8371

Helena Silva

Professora Associada

Universidade de Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal

Centro de Investigação e Intervenção Educativas (CIIE), Universidade do

Porto, Portugal

Email: helsilva@utad.pt

ORCID: 0000-0003-2855-9634

Eva Morais

Professora Auxiliar
Universidade de Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal
Polo do Centro de Matemática da Universidade do Minho (CMAT-UTAD), Vila
Real, Portugal
Email: emorais@utad.pt
ORCID: 0000-0002-3815-9821

Correspondência

José Lopes
Professor Associado
Escola de Ciências Humanas e Sociais
Universidade de Trás-os-Montes e Alto Douro
Quinta de Prados
5000-801 Vila Real

Data de submissão: Setembro 2018

Data de avaliação: Dezembro 2018

Data de publicação: Julho 2019