

Escala Informatizada de Avaliação das Características Criativas: Evidências de Validade de Conteúdo

Talita Fernanda da Silva¹, Tatiana de Cássia Nakano
Pontifícia Universidade Católica de Campinas, Campinas-SP, Brasil

RESUMO

O uso da tecnologia tem sido recomendado com um recurso adicional na área da avaliação psicológica. Dentro dessa proposta, a Escala Informatizada de Avaliação das Características Criativas foi desenvolvida. Com o objetivo de verificar a adequação de seus itens, dois estudos foram conduzidos. No primeiro, cinco estudantes de pós-graduação atuaram como juízes, classificando os 64 itens do instrumento em oito características (Fluência, Flexibilidade, Elaboração, Originalidade, Expressão de Emoção, Fantasia, Perspectiva Incomum, Analogias/Metáforas). Os resultados demonstraram a adequação dos itens, com a maioria deles apresentando índices de concordância acima de 80%. Também foram encontrados valores adequados de Kappa para os juízes. Na análise 2, 50 estudantes, com idades entre 18 a 50 anos de idade ($M=32,7$; $DP=10,27$) participaram de estudo piloto, auxiliando na reformulação e adequação de diferentes aspectos da escala. Os resultados apontaram adequação da escala ao construto e à população que pretende avaliar.

Palavras-chave: avaliação psicológica; avaliação psicológica informatizada; criatividade.

ABSTRACT – Computerized Scale of Creative Characteristics: Evidence of Content Validity

The use of technology has been recommended with an additional feature in the area of psychological assessment. Within this proposal, the Computerized Scale of Creative Characteristics was developed. In order to verify the adequacy of their items, two studies were conducted. In the first one, five graduate students acted as judges, classifying the 64 items of the instrument into eight characteristics (Fluency, Flexibility, Elaboration, Originality, Expression of Emotion, Fantasy, Unusual Perspective, Analogies / Metaphors). The results demonstrated the adequacy of the items, with most of them presenting agreement indexes above 80%. Adequate Kappa values were also found for the judges. In the study 2, 50 students, aged between 18 and 50 years ($M=32.7$; $SD=10.27$) participated in a pilot study, helping to reformulate and adapt different aspects of the scale. The results pointed to the adequacy of the scale to the construct and to the population that it intends to evaluate.

Keywords: psychological assessment; computerized assessment; creativity.

RESUMEN – Escala Informatizada de Evaluación de las Características Creativas: evidencias de validez de contenido

El uso de la tecnología está siendo recomendada como un recurso adicional en el área de la evaluación psicológica. Dentro de esta propuesta, la Escala Informatizada de Evaluación de las Características Creativas fue desarrollada. Con el objetivo de verificar la adecuación de sus ítems, se realizaron dos estudios. En el primero, cinco estudiantes de posgrado actuaron como jueces, clasificando los 64 ítems del instrumento en ocho características (Fluencia, Flexibilidad, Elaboración, Originalidad, Expresión de Emoción, Fantasía, Perspectiva Inusual, Analogías / Metáforas). Los resultados demostraron la adecuación de los ítems, con la mayoría de ellos presentados índices de concordancia por encima del 80%. También fueron encontrados valores adecuados de Kappa para los jueces. En el segundo estudio, participaron 50 estudiantes, con edades entre 18 a 50 años de edad ($M=32,7$, $SE=10,27$), en un estudio piloto, auxiliando en la reformulación y adecuación de diferentes aspectos de la escala. Los resultados apuntaron la adecuación de la escala al constructo y a la población que pretende evaluar.

Palabras clave: evaluación psicológica; evaluación informatizada; creatividad.

O avanço da tecnologia tem sido cada vez mais presente no cotidiano das pessoas, nos mais diferentes contextos, tais como trabalho, escola, na execução de diferentes tarefas, e inclusive, nos momentos de lazer com a família e/ou com os amigos (Katsurayama, Silva, Eufrazio, Souza, & Becker, 2012; Olea, Ponsoda, & Prieto, 2006). Por tal motivo, atualmente muito tem se buscado compreender a

respeito do impacto da tecnologia na vida das pessoas, de modo que essa temática tem ocupado um papel de destaque em estudos recentes no Brasil e em outros países (Almerich et al., 2005; Baracho & Freitas Junior, 2019; Prado, 2005; Olea, Abad, Ponsoda, & Ximénez, 2004).

Diante desse cenário, a Psicologia tem buscado incorporar, mais recentemente, a informática e a

¹ Endereço para correspondência: Avenida Palmeira Real, 705, Alto das Palmeiras, 13301-310, Itu, SP. E-mail: talita398@gmail.com
A autora agradece a Capes pela concessão de bolsa de auxílio à pesquisa.

informatização dos dados na sua prática profissional (Katsurayama et al., 2012), mais especificamente no campo da avaliação psicológica (Butcher, Perry, & Atlis, 2000; Butcher, Perry, & Hahn, 2004; Joly, Martins, Abreu, Souza, & Cozza, 2004; Joly & Reppold, 2010; Olea, Abad, & Barrada, 2010). No entanto, ainda que o caminho do desenvolvimento da avaliação informatizada venha se mostrando promissor (Nakano, 2013), visto que os recursos da tecnologia da informatização muito poderiam contribuir para os avanços dos recursos que se necessita quando se trata da precisão e fidedignidade dos instrumentos psicológicos, essa tecnologia ainda se mostra pouco explorada. Tal fato pode ser constatado, por meio de consulta, em março de 2019, ao Sistema de Avaliação dos Testes Psicológicos. Na ocasião, foi possível identificar a presença de apenas dois instrumentos informatizados, sendo: Escala de Matrizes de Vienna – 2 Versão Informatizada (WMT-2) (Formann, Waldherr, & Piswanger, 2017) e o Sistema Multimídia de Habilidades Sociais de Crianças (Del Prette & Del Prette, 2005).

Os testes informatizados têm sido definidos a partir de uma variedade de termos e conceitos, como testes baseados, administrados, adaptados e avaliados em computadores, avaliação informatizada ou assistida em teste informatizado, segundo Devriendt (2008) e a *International Testing Commission* (2000, 2005, 2009, 2010). A diferença básica, desse tipo de teste e aquele aplicado via tradicional (por meio de lápis e papel), consiste no fato de que o segundo tipo realiza a avaliação das capacidades/habilidades por meio do emprego de um software e hardware usando o dispositivo da tela do computador. Assim, um teste pode ser chamado e considerado informatizado quando pode ser organizado e aplicado via computador e/ou internet (Andriola, 2003; Joly & Noronha, 2006).

Torna-se importante ressaltar que tal característica não o diferencia, em nada, da necessidade de apresentação de estudos voltados à investigação de suas qualidades psicométricas, as mesmas exigidas em relação aos instrumentos do tipo lápis e papel, conforme Resolução 09/18 (CFP, 2018). As diferenças se situam, segundo Olea, Ponsoda e Prieto (2006) na necessidade de verificação de uma série de requisitos, os quais envolvem os aspectos técnicos e tecnológicos presentes, a qualidade do teste informatizado, o formato de disponibilização *off-line* e/ou *on-line*, e assim os níveis de segurança e privacidade, ressaltados pela *International Testing Commission* (2005).

Perante a constatação da existência de um número ainda restrito de instrumentos psicológicos que se caracterizam como informatizados, das vantagens que a utilização das tecnologias informatizadas na área da avaliação psicológica, bem como lacunas existentes em relação ao número de instrumentos disponíveis no Brasil para avaliação da criatividade, o processo de desenvolvimento de um teste informatizado para avaliação de características criativas foi iniciado, intitulado Escala Informatizada

de Avaliação das Características Criativas. Importante destacar que, quando se trata do construto criatividade, nenhum instrumento de avaliação que faça uso do recurso da informática, em nenhuma de suas etapas, tanto em termos de aplicação, correção ou avaliação, é encontrado no Brasil.

Embasado em uma compreensão multidimensional da criatividade, a qual envolve uma série de aspectos, tais como a pessoa, produto, processo e ambiente criativo, o instrumento se foca no primeiro aspecto e contempla a avaliação de uma série de características associadas à criatividade, as quais vêm sendo ressaltadas na literatura científica nacional e internacional. Maiores detalhes acerca do instrumento serão apresentados por ocasião da sua descrição no tópico referente à metodologia. Perante a necessidade da investigação de suas qualidades psicométricas, aqui são relatados dois estudos, relacionados à busca por evidências de validade do instrumento.

Método

Estudo 1 – Busca por Evidências de Validade de Conteúdo

O objetivo desse estudo foi verificar evidências de validade de conteúdo, mais especificamente, a adequação dos itens que compõem a Escala de Avaliação das Características Criativas, por meio da análise da concordância de juízes. Esse tipo específico de evidência de validade, considerado um dos métodos a serem empregados durante o processo inicial de construção de instrumento, antes da sua aplicação empírica (Alexandre & Colucci, 2011), visou investigar a clareza, representatividade e relevância dos itens desenvolvidos.

Participantes

A amostra foi composta por cinco juízes independentes, estudantes de um programa de pós-graduação em Psicologia, de uma universidade privada situada no interior do Estado de São Paulo, sendo eles vinculados à linha de pesquisa “Instrumentos e Processos em Avaliação Psicológica”. Tal procedimento foi utilizado como forma de garantir que todos possuíssem conhecimento da área de Avaliação Psicológica e de construção de instrumentos. Destes, quatro eram estudantes de doutorado, e um estudante de mestrado, sendo quatro do sexo feminino.

Instrumento

Escala de Avaliação das Características Criativas – versão lápis e papel. Nesse estudo, o instrumento foi trabalhado na versão lápis e papel, composta por 64 itens, sendo oito para cada característica criativa avaliada: Fluência (capacidade de fornecer grande quantidade de respostas), Flexibilidade (pensar em diferentes categorias/possibilidades de respostas), Elaboração (habilidade de adicionar detalhes à solução

proposta), Originalidade (criação de respostas incomuns), Expressão de Emoção (inserção de emoções no processo de busca por uma resposta), Fantasia (presença de seres imaginários, ficção científica ou conto de fadas), Perspectiva Incomum (ver as situações sobre uma perspectiva não usual), Uso de Analogias/Metáforas (realização de associações entre ideias).

Procedimentos

A coleta de dados para a realização deste estudo foi iniciada após aprovação da pesquisa pelo Comitê de Ética (CAAE: 60504116.2.0000.5481). Para coleta de dados, a pesquisadora desenvolveu um documento do tipo formulário, o qual foi entregue aos juízes com explicações acerca de como deveria ser realizada a atividade de avaliação. O documento apresentava as definições de todas as características a serem avaliadas no instrumento, bem como uma lista contendo os 64 itens a serem avaliados. O trabalho dos juízes consistiu em ler cada uma das frases, analisando e julgando qual das características criativas ele acreditava que o conteúdo da frase abordava.

Com a finalidade de evitar que os itens estivessem agrupados de acordo com a característica que avalia, eles foram organizados em uma única lista, de maneira aleatória, com o intuito de dificultar, aos juízes, a descoberta dos itens, verificando a disposição deles, de forma que não fosse possível identificar os subgrupos de itens por similaridade ou proximidade, conforme recomendado por Pasquali (2010). Após o recebimento de todos os formulários preenchidos, a pesquisadora efetuou o levantamento das categorias apontadas, pelos juízes, para cada um dos itens. Para a análise dos dados foi realizado o uso de dois procedimentos, sendo a estimativa do Índice de Concordância (IC) para cada item da escala, e posteriormente o cálculo do coeficiente de Kappa, para cada juiz participante, sendo tais métodos recomendados como complementares (Perroca & Gaidzinski, 2003).

Análise de dados

No primeiro procedimento foi calculada a porcentagem de concordância entre os juízes, para cada item.

Aqueles itens que obtiverem concordância acima de 80% foram considerados adequados, sendo selecionados para compor a escala. Os demais, que ficaram abaixo dessa porcentagem, foram redigidos novamente ou excluídos da versão final do instrumento, havendo ainda a possibilidade de que algum item fosse realocado em outra categoria, diferente daquela para a qual foi originalmente desenvolvido, caso houvesse concordância dos juízes (acima de 80%) nessa nova categoria.

Em relação ao segundo procedimento de análise, a estimativa do coeficiente Kappa, tem se mostrado um método bastante utilizado (Miura, Gallani, Domingues, Rodrigues, & Stoller, 2010) quando vários avaliadores categorizam cada grupo de objetos ou sujeitos em categorias nominais (Alexandre & Coluci, 2011). Visa descrever e testar o grau de concordância (confiabilidade e precisão) na classificação de diferentes juízes (Perroca & Gaidzinski, 2003). Para a interpretação desse coeficiente, os valores de Kappa, fornecidos por Landis e Koch (1977) foram utilizados: concordância quase perfeita (entre 0,80 e 1,00), concordância substancial (0,60 a 0,80), concordância moderada (0,40 a 0,60), concordância regular (0,20 a 0,40) e concordância discreta (0,00 a 0,20).

Resultados

A primeira análise realizada utilizou o índice de concordância dos juízes (IC), por meio da estimativa da porcentagem. Importante destacar que os itens que não apresentaram porcentagem de concordância ideal (acima de 80%) foram destacados em cinza na Tabela, para melhor visualização.

Na Tabela 1, é possível notar que, dos 64 itens analisados, a maior parte deles, 81,25% apresentou valores interpretados como concordância quase perfeita (índices maiores de 80%), mostrando-se adequados e, consequentemente, selecionados para a primeira versão do instrumento. Também é possível verificar que as características de expressão de emoção e fantasia apresentaram concordância entre juízes adequada para todos os seus itens, de modo que todos foram mantidos.

Tabela 1
Índice de Concordância Entre Juízes para os Itens da Escala

Característica	Item	Concordância	Característica	Item	Concordância
Fluência	1	100	Elaboração	17	100
	2	60		18	100
	3	60		19	100
	4	80		20	100
	5	60		21	100
	6	80		22	100
	7	80		23	60
	8	80		24	80

Tabela 1 (continuação)
Índice de Concordância Entre Juízes para os Itens da Escala

Característica	Item	Concordância	Característica	Item	Concordância
Flexibilidade	9	80	Originalidade	25	80
	10	60		26	80
	11	80		27	40
	12	100		28	100
	13	60		29	80
	14	80		30	80
	15	60		31	100
	16	100		32	80
Expressão de Emoção	33	80	Fantasia	41	80
	34	100		42	80
	35	100		43	100
	36	80		44	80
	37	100		45	100
	38	80		46	100
	39	100		47	100
	40	100		48	100
Perspectiva Incomum	49	80	Analogias e Metáforas	57	80
	50	60		58	100
	51	80		59	80
	52	80		60	100
	53	40		61	80
	54	80		62	80
	55	100		63	80
	56	60		64	40

Outros 12 itens não apresentaram concordância adequada entre os juízes. Três pertenciam à característica de fluência (itens 2, 3 e 5), três avaliavam a flexibilidade (itens 10, 13 e 15), três eram referentes à perspectiva incomum (itens 50, 53 e 56), um de elaboração (item 23), um de originalidade (item 27) e um de analogias/metáforas (item 64). Destes, três obtiveram somente 40% de concordância e nove tiveram 60% de concordância, sendo que todos os itens nessa condição foram reformulados, recebendo nova redação ou novo conteúdo.

Importante destacar que algumas situações não previstas inicialmente foram verificadas após o julgamento dos juízes. O item 10 indicou concordância substancial (entre 60% e 80%) para a característica Fluência e não para a característica criativa Flexibilidade, para a qual originalmente foi desenvolvido. Por tal motivo, as autoras decidiram pela sua exclusão e desenvolvimento de um novo item para substituí-lo. Situação semelhante foi verificada em relação ao item 16, o qual apresentou concordância de 100% para a característica Fluência e não Flexibilidade. Por tal motivo, ele também foi excluído, tendo-se criado um novo item para Flexibilidade. O item 29 apresentou concordância de 80% entre os juízes para a característica de Fluência e não para Originalidade, sua categoria original. Diante

da análise de seu conteúdo, conduzida pelas autoras, o mesmo foi realocado para a característica de Fluência, na qual obteve concordância dos juízes.

Posteriormente uma nova rodada de avaliação pelos juízes foi realizada com os itens que inicialmente não apresentaram concordância adequada. Os resultados apresentados pelos novos itens são fornecidos na Tabela 2. Importante salientar que nessa segunda rodada de avaliação, os juízes receberam somente os novos itens para julgamento ($n=12$).

Na Tabela 2, é possível verificar que todos os novos itens de Fluência, Elaboração e Originalidade testados apresentaram concordância maior do que 80%, mostrando-se adequados nessa nova versão. Em outro oposto, todos os itens de Flexibilidade, Perspectiva Incomum e Uso de Analogias e Metáforas ainda continuaram apresentando concordância baixa entre os juízes. Com isso, pode-se verificar que, dos 12 itens reformulados, sete ainda apresentaram ausência de concordância na avaliação dos juízes. Tais itens foram, mais uma vez, reescritos e passaram por uma terceira e última rodada de avaliação, sendo importante destacar que, em alguns casos, mais de uma versão do mesmo item foi desenvolvida, visando-se verificar se alguma delas apresentava melhor resultado.

Tabela 2
Resultados da Segunda e Terceira Rodada de Análise dos Índices de Concordância Entre Juízes

Segunda rodada de análise					
Característica	Item	Concordância	Característica	Item	Concordância
Fluência	2	100	Elaboração	23	100
	3	80	Originalidade	27	80
	5	100		29	100
Flexibilidade	10	60	Perspectiva Incomum	50	60
	15	60		53	60
	16	20		56	60
Analogias	63	60			
Terceira rodada de análise					
Característica	Item	Concordância	Característica	Item	Concordância
Flexibilidade	10	60	Perspectiva Incomum	50	80
	10	60		50	100
	15	20		53	100
	15	60		53	80
	16	60		56	80
	16	60		56	80
Analogias	63	100			
	63	60			
	63	100			

Após terceira rodada de análise dos itens, os seis itens construídos para avaliação da Perspectiva Incomum apresentaram concordância adequada, o mesmo acontecendo em relação à dois dos três itens analisados na característica de Analogias e Metáforas. Novamente verificou-se dificuldade dos juízes em compreender a definição de Flexibilidade e julgar seus itens visto que, dos seis novos itens, todos apresentaram resultado abaixo da porcentagem de concordância esperada (80%).

Diante desse fato e da importância dessa característica para a expressão criativa, uma hipótese de que a definição dessa característica, apresentada aos juízes não estivesse clara foi pensada. Assim, uma nova definição para tal característica criativa foi elaborada, acompanhada da elaboração de novos itens ($n = 11$). Ao final de nova rodada de avaliação, três apresentaram concordância ideal e foram selecionados para compor a primeira versão do instrumento, os quais foram somados aos cinco que se mostraram adequados na primeira rodada de avaliação. Finalmente o procedimento de análise dos índices de concordância estava finalizado, tendo-se oito itens para cada uma das oito características criativas avaliadas pelo instrumento.

Em seguida, o segundo procedimento de análise, a estimativa do coeficiente Kappa, foi empregado, para descrever e testar o grau de concordância na classificação de diferentes juízes. Os resultados foram analisados em relação ao número de itens classificados em cada área, número de acertos, porcentagem de acertos e valor do Kappa, sendo apresentados na Tabela 3.

A partir da Tabela 3, verifica-se que, dos cinco juízes, quatro apresentaram valor de Kappa desejado, ou seja, acima de 0,80. A exceção ocorre em relação ao juiz 4. Na Tabela 3, ainda é possível observar uma tendência dos juízes em classificar mais itens na característica criativa C7=Perspectiva Incomum ($n=46$) e menos na característica C8=Analogias e Metáforas ($n=36$). No entanto, quando se analisa a média de porcentagem de acertos por característica, o que se nota é que os juízes tiveram mais facilidade em identificar os itens pertencentes às características de Elaboração (acertando, em média, 97,5% dos itens pertencentes a essa característica), seguida pelas de Expressão da Emoção e Fantasia (92,5% de acertos em cada). Por outro lado, as características que se mostraram mais difíceis foram as de Flexibilidade e Originalidade, com 82,5% e 85,0% de acertos, respectivamente.

Algumas hipóteses para tais resultados envolvem a possibilidade de que, as características cujas médias de acerto se mostraram mais altas apresentaram definição mais precisa e clara para os juízes, de modo que tal fato facilitou a identificação de seus itens ou que os itens representassem, mais diretamente, em seu conteúdo, aspectos presentes nas definições fornecidas no formulário dos juízes. A situação oposta pode ser hipotetizada em relação às características em que os juízes tiveram mais dificuldade de identificar seus itens. Especificamente em relação à Flexibilidade, a dificuldade dos juízes ficou explicitada perante a necessidade, já relatada anteriormente, de uma rodada extra de avaliação, contendo somente novos itens pertencentes a esta característica.

Tabela 3
Estatística Kappa Para Avaliação dos Juizes para Cada Uma das Características Criativas

Juiz	Kappa	Itens Classificados	C1	C2	C3	C4	C5	C6	C7	C8
1	0,94	Nº de itens classificados em cada área	8	9	8	8	8	8	8	8
		Acertos	7	7	8	8	8	8	8	8
	0,001	% de acertos	87,5	87,5	100	100	100	100	100	100
2	0,91	Nº de itens classificados em cada área	6	11	9	7	7	8	8	8
		Acertos	6	8	8	7	7	8	7	8
	0,001	% de acertos	75	100	100	87,5	87,5	100	87,5	100
3	0,89	Nº de itens classificados em cada área	6	10	9	6	9	7	9	8
		Acertos	6	8	8	6	8	7	7	8
	0,001	% de acertos	75	100	100	75	100	87,5	87,5	100
4	0,73	Nº de itens classificados em cada área	9	8	8	12	6	7	10	4
		Acertos	8	6	7	7	6	6	6	3
	0,001	% de acertos	100	75	87,5	87,5	75	75	75	37,5
5	0,89	Nº de itens classificados em cada área	10	4	9	6	8	8	11	8
		Acertos	8	4	8	6	8	8	8	8
	0,001	% de acertos	100	50	100	75	100	100	100	100
Média da porcentagem de acertos			87,5	82,5	97,5	85	92,5	92,5	90	87,5
Total			39	42	43	39	38	38	46	36

Nota. C1=Fluência; C2=Flexibilidade; C3=Elaboração; C4=Originalidade; C5=Expressão de Emoção; C6=Fantasia; C7=Perspectiva Incomum; C8=Analogias e Metáforas

Discussão

O estudo 1 teve, como objetivo, a busca por evidências de validade de conteúdo de uma escala em processo de desenvolvimento. Por meio do uso de dois diferentes métodos (porcentagem de concordância e coeficiente Kappa), os itens puderam ser reformulados de modo a apresentarem-se de forma clara, representativa e relevante para o construto que se propõem a avaliar, segundo julgamento de juizes. Tal etapa se mostrou fundamental, principalmente, por auxiliar as pesquisadoras na seleção e manutenção de itens que se mostraram adequados, bem como na reformulação de outros que não estavam sendo corretamente interpretados. Igualmente, possibilitou a reescrita da definição da característica de flexibilidade, antes que qualquer procedimento empírico fosse conduzido com o instrumental.

Estudo 2 – Estudo Piloto de Aplicação da Escala Informatizada de Avaliação das Características Criativas

O segundo estudo conduzido envolveu a aplicação dos itens anteriormente selecionados após análise de juizes em uma amostra reduzida de sujeitos, com o objetivo de verificar sua adequação à população-alvo, dentro de um procedimento chamado de estudo piloto ou grupo focal. Almejou-se avaliar o quanto os itens são compreensíveis e claros para todos os membros da população a qual o instrumento se destina (Pasquali, 2010).

Uma das formas de se realizar esse tipo de estudo consiste em aplicar o instrumento em uma amostra de

sujeitos com características similares à população para a qual o instrumento será destinado. Nesse processo, o autor salienta que tal procedimento tem, como finalidade, levantar, junto aos participantes, eventuais dúvidas que os itens suscitarem. A esse grupo apresenta-se item por item, pedindo que seu conteúdo seja explicado pelos membros do grupo, solicitando aos participantes que verbalizem o que entenderam do item, o que ele pergunta, bem como verificar se todos o compreendem da mesma forma (Pacico, 2015).

Se, durante esse processo, a reprodução do item não deixar nenhuma dúvida, pode-se supor que o item é corretamente compreendido. No entanto, se surgirem divergências na reprodução do item ou se o pesquisador perceber que está sendo entendido diferentemente do que ele julga que deveria ser entendido, esse item tem problemas, devendo ser modificado e passar, novamente, por outro grupo piloto após sua reformulação ou, ainda, ser eliminado. O pesquisador pode também explicar, ao grupo, o que ele pretendia dizer com tal item, sendo solicitado aos próprios sujeitos que ofereçam sugestões sobre como este deveria ser formulado (Pasquali, 2010). Tal método foi adotado no presente estudo, descrito a seguir.

Participantes

A amostra foi composta por 50 participantes, na faixa etária de 18 a 50 anos ($M=32,7$; $DP=10,24$), de ambos os sexos (sendo 43 do sexo feminino), selecionados por conveniência dentre os estudantes de pós-graduação *lato sensu* de um curso de Psicologia Organizacional e do

Trabalho oferecido em uma faculdade particular localizada no interior do estado de São Paulo.

Instrumento

Escala Informatizada de Avaliação das Características Criativas – Versão Informatizada. A versão da escala utilizada nesse estudo foi aquela composta por 64 itens, após os resultados do Estudo 1, sendo que, no estudo aqui relatado, o instrumento já foi utilizado em sua versão informatizada.

Procedimentos

O Estudo 2 ocorreu após a informatização da escala, realizada com auxílio de um profissional da área de Informática. Durante esse processo, uma série de cuidados e decisões foram tomadas em relação à seleção do melhor formato de apresentação dos itens, do sistema de acesso, da inserção do TCLE, dos dados sociodemográficos e de identificação dos sujeitos, formato do banco de dados gerado, bem como a disponibilização e testagem do funcionamento do sistema.

Em dia e horários previamente agendados com os professores e alunos, a coleta de dados aconteceu em uma sala de informática da própria universidade. A aplicação foi realizada em uma única sessão, com duração estimada de 40 minutos, de forma coletiva. Vale destacar que a aplicação foi dividida em dois grupos, sendo o primeiro composto por 35 estudantes, e posteriormente um grupo de 15 estudantes, visando que condições adequadas de testagem fossem garantidas.

Após orientações iniciais, solicitou-se que cada estudante ligasse seu computador e acessasse o site www.escaladecriatividade.com.br. Nessa primeira parte, de 35 computadores, apenas dois não conseguiram acessar diretamente o site pelo *Google Chrome*, mas foi possível acessá-lo pelo navegador *Internet Explorer*. Antes que os participantes iniciassem o processo de resposta, a pesquisadora explicou as etapas que eles iriam encontrar no site.

Inicialmente, o Termo de Consentimento Livre e Esclarecido, contendo todos os dados da pesquisa, como informações sobre sua participação, garantia de sigilo e anonimato, assim como os riscos e benefícios advindos da sua participação seria encontrado no endereço fornecido. O participante deveria clicar na opção: “aceito participar da pesquisa”, para que pudessem acessar a segunda parte e preencher alguns dados pessoais e sociodemográficos (*e-mail*, idade, gênero, estado civil, cidade de moradia, estado, escolaridade).

Após o preenchimento dos dados solicitados, na próxima página, tiveram acesso à escala na tela do computador. Inicialmente apareceram para o participante, instruções sobre como responder à escala, assim como dois exemplos de como efetuar o preenchimento da sua resposta. Na parte das instruções, apenas três participantes tiveram dúvidas sobre a compreensão da escala Likert, sendo estas

esclarecidas pela pesquisadora antes do início do processo de resposta. Durante a atividade eram apresentados oito itens por vez (e por página). Tendo respondido a todos os itens da página, automaticamente o participante era direcionado à próxima, contendo os próximos oito itens e assim sucessivamente até que tivesse completado o processo de resposta dos 64 itens da escala. O participante conseguia retornar às páginas já respondidas, apenas não conseguia enviar a pesquisa se não tivesse respondido todos os itens.

A tarefa consistiu na leitura e avaliação do conteúdo presente em cada item, julgando o quanto ele se identificava. O participante deveria selecionar, dentro de uma escala Likert de cinco pontos (sendo: descreve-me mal, descreve-me pouco, descreve-me, descreve-me bem, descreve-me muito bem) a resposta que melhor se adequasse à intensidade com que o conteúdo o representava. Em um segundo momento, cada participante recebeu um papel para que pudessem anotar suas dúvidas, queixas e sugestões a respeito da escala, podendo realizar essa tarefa durante e/ou após a participação, de acordo com sua preferência. Concomitantemente à aplicação, a pesquisadora anotou todos os comentários, dificuldades e comportamentos apresentados pelos participantes.

No segundo grupo, apenas um estudante teve dúvidas quanto às instruções gerais, e principalmente com a escala Likert, sendo que todas as suas dúvidas foram esclarecidas. Ao término da aplicação, todos os estudantes foram convidados para retornarem na sala de informática novamente ($n=50$), sendo convidados a apresentarem suas impressões sobre a escala. Dentre os principais comentários, o desejo de ter recebido um resultado após responder a pesquisa foi o mais enfatizado. O motivo disso não ter sido possível foi esclarecido, argumentando-se sobre o fato de que o instrumento ainda se encontrava em fase de investigação. No entanto, logo após sua participação, os estudantes puderam participar de uma palestra sobre criatividade. A pesquisadora, após a aplicação do estudo piloto, leu todas as fichas com anotações de sugestões dos participantes, com o objetivo de verificar os apontamentos realizados. Os resultados são apresentados a seguir.

Resultados

Em relação às anotações dos 50 participantes, 26 deles fizeram registros na folha fornecida. Dentre os principais pontos positivos apontados, podem ser citados alguns comentários, tais como: “fácil acesso ao teste”, “teste rápido e prático”, “proposta boa para avaliar pessoas”, “teste interessante”, “teste que pode ter uma boa aceitação do respondente”, “bem elaborado”, “pesquisa inovadora para o mercado”, “site bem feito”, entre outros.

Uma série de outros pontos, citados como negativos pelos estudantes, e que poderiam ser alvo de melhoria, envolveram: “resultado deveria sair após finalizar o teste”, “letra das questões poderia ser maior”, “a tela em que ficam as perguntas poderia ser maior”, “a palavra implícita e maleável

deveria ser trocada, porque tem pessoas que não sabem seu significado", "inserir imagens junto com as perguntas", "fiquei confusa com a escala na hora de responder, não sabia identificar o grau de 1 a 5, tive que ler algumas vezes".

Interessantemente, um comentário que foi bastante citado, mas que não mostrou consenso entre os participantes, ora sendo citado como positivo e, em outras opiniões, como ponto frágil foi o fato de que, na visão de alguns participantes, a escala era grande e poderia ser menor, já na visão de outros a escala era pequena e poderia ser colocado mais itens. Ainda, durante a aplicação do estudo, a pesquisadora realizou observações e anotações e verificou a necessidade de melhoria na explicação acerca da escala Likert, visto que quatro, de 50 participantes, apresentaram dificuldades em compreendê-la.

Como resultado do estudo piloto conduzido, proveniente das anotações fornecidas pelos estudantes e registros de observações da pesquisadora, uma série de ajustes foram realizados na escala, os quais envolveram: 1. tamanho da letra para escrita dos itens (ampliando-a); 2. aumento do tamanho da tela em que os itens estavam inseridos; 3. ampliação e revisão das explicações da escala Likert nas instruções da escala; 4. reescrita das instruções, visando-se melhor organização; 5. as palavras "implícita" e "maleável" foram removidas e substituídas por outras palavras nos itens em que elas apareciam; 6. uma cartilha sobre criatividade foi desenvolvida para ser distribuída *on-line* a todos participantes da pesquisa.

Discussão

Como resultado, uma série de ajustes foi realizada na escala, desde aspectos que envolviam o *layout* da escala informatizada, para uma melhor interação do sujeito com a escala (tamanho da letra, distribuição de itens por tela, gráfico representando a escala Likert para facilitação da leitura e compreensão do sujeito, cores, entre outros), até mesmo a revisão das instruções da escala, substituição de palavras dentro do item para melhor entendimento, reformulação de itens, desenvolvimento de um tipo de devolutiva ao participante, entre outras.

Em relação aos itens, vale destacar que, ainda que comentários dos participantes sobre a similaridade entre alguns deles tenham sido notados, nenhuma alteração foi feita nesse sentido, visto que tal situação decorre, na realidade, da similaridade encontrada entre algumas definições de características criativas selecionadas para compor o instrumento. Pode-se citar, por exemplo, a proximidade entre as características de fluência, flexibilidade e originalidade,

já bastante discutida na literatura científica, notadamente a internacional (Chase, 1985; Heausler & Thomson, 1988; Nakano & Primi, 2012; Runco & Mraz, 1992).

A independência das características avaliadas tem sido questionada, diante da constatação de alta correlação e, conseqüentemente, possível redundância encontrada entre algumas das características avaliadas na maior parte dos instrumentos de avaliação da criatividade (Nakano & Primi, 2012). Tal situação poderá, ou não, ser notada no instrumento em processo de desenvolvimento, mas somente após a condução de estudos voltados à sua estrutura fatorial, análise que permitirá identificar a distinção e independência entre as características criativas que a escala se propõe a avaliar. Caso tal situação não seja verificada empiricamente, medidas voltadas ao controle dessa condição serão pensadas.

Considerações Finais

Diante do exposto, foi possível perceber que o estudo piloto se mostrou de extrema relevância e atingiu seu objetivo de verificar a adequação dos itens e funcionalidade junto à população-alvo, bem como possibilitar ajustes voltados à sua melhor compreensão e aplicabilidade. A partir deste estudo também foi possível perceber que o sistema de envio das respostas do participante ao aplicador teve um bom funcionamento, comprovado por meio do recebimento das respostas dos participantes, de acordo com o formato desejado para o banco de dados. Com isso, o estudo-piloto se mostrou de extrema relevância e atingiu seu objetivo de verificar a adequação dos itens e funcionalidade junto à população-alvo, permitindo mudanças voltadas à melhoria na escala em desenvolvimento, possibilitando, ainda, verificar, na prática, o funcionamento da informatização da escala.

Os resultados dos dois estudos conduzidos apontaram para evidências positivas e favoráveis ao instrumento, envolvendo aspectos relacionados à adequação dos itens ao conteúdo proposto, assim como melhorias na sua estrutura enquanto instrumento, envolvendo desde as instruções do teste, funcionalidade, até mesmo, estética. Considerando-se que a busca por evidências de validade se mostra um processo que deve ser contínuo e cumulativo, mostra-se necessária a condução de outros estudos até que uma quantidade mínima de propriedades psicométricas seja evidenciada. Dentre esses, destaca-se a necessidade de busca por outros tipos de evidências de validade e estudos voltados à precisão e normatização, a fim de que possa ser disponibilizado, futuramente, para uso profissional.

Referências

- Alexandre, N. M. C., & Coluci, M. Z. O. (2011). Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Ciência e Saúde Coletiva*, 16(7), 3061-3068. doi: 10.1590/S1413- 81232011000800006

- Almerich, G., Suárez, J. M., Orellana, N., Belloch, C. B. R., Bo, R., & Gastaldo, I. (2005). Diferencias en los conocimientos de los recursos tecnológicos en profesores a partir del género, edad y tipo de centro. *Revista Electrónica de Investigación y Evaluación Educativa*, 11(2), 127-146. Recuperado de <http://www.redalyc.org/articulo.oa?id=91611203>
- Andriola, W. B. (2003). Uso de computadores na avaliação psicológica: Estudo de sua influência sobre o desempenho individual em um teste de raciocínio numérico (RN). *Interações*, 3(15), 105-124. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1413-29072003000100006
- Baracho, R. M. A., & Freitas Junior, C. A. (2019). Educação 3.0: A educação da 4ª Revolução Industrial. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 14(1), 22-30. doi: 10.22478/ufpb.1981-0695.2019v14n1.44468
- Butcher, J. N., Perry, J. N., & Atlas, M.M. (2000). Validity and utility of computer-based test interpretation. *Psychological Assessment*, 12(1), 6-18. doi: 10.1037/1040-3590.12.1.6
- Butcher, J. N., Perry, J., & Hahn, J. (2004). Computers in clinical assessment: Historical developments, present status, and future challenges. *Journal of Clinical Psychology*, 69(3), 331-345. doi: 10.1002/jclp.10267
- Chase, C. I. (1985). Review of the Torrance Tests of Creative Thinking. Em J. V. Mitchell Jr. (Ed.), *The ninth mental measurements yearbook* (pp. 1631-1632). Lincoln: University of Nebraska Press.
- Del Prette, Z. A. P., & Del Prette, A. (2005). *Sistema multimídia de habilidades sociais de crianças*. São Paulo: Casa do Psicólogo.
- Devriendt, Y. A. (2008). Computer-based testing. Em M. Born, C. D., Foxcroft, & R. Butter (Eds.), *Online Readings in Testing and Assessment*. Recuperado de <http://www.intestcom.org/Publications/ORTA.php>
- Formann, A. K., Waldherr, K., & Piswanger, K. (2017). *Teste de Matrizes de Viena*. São Paulo: Hogrefe.
- Heausler, N. L., & Thompson, B. (1988). Structure of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*, 48, 463-468. doi: 10.1177/0013164488482021
- International Test Commission (2000). *ITC International Test Commission*. Recuperado de <http://www.intestcom.org/>
- International Test Commission (2005). *International Test Commission*. Guidelines on computer-based and internet-delivered testing. Recuperado de <http://www.intestcom.org/>
- International Test Commission (2009). *International Test Commission. A test-takes guide to technology based testing*. Recuperado de <http://www.intestcom.org/>
- International Test Commission (2010). *International Test Commission*. Recuperado de <http://www.intestcom.org/>
- Joly, M. C. R. A., & Noronha, A. P. P. (2006). Reflexões sobre a construção de instrumentos psicológicos informatizados. Em A. P. P. Noronha, A. A. Santos, & F. F. Sisto (Eds.), *Facetas do Fazer em Avaliação Psicológica* (pp.95-105). São Paulo: Vetor.
- Joly, M. C. R. A., & Reppold, C. T. (2010). *Testes informatizados para a avaliação psicológica e educacional*. E-book. São Paulo: Casa do Psicólogo.
- Joly, M. C. R. A., Martins R. X., Abreu, M. C., Souza, P. R. R., & Cozza H. F. P. (2004). Análise da produção científica em avaliação psicológica informatizada. *Avaliação Psicológica*, 3(2), 121-129. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712004000200007
- Katsurayama, M. Silva, S. R., Eufrazio, W. N., Souza, R. S. A., & Becker, M. A. A. (2012). Testes informatizados como auxílio na seleção em recursos humanos. *Psicologia: teoria e prática*, 14(2), 141-151. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1516-36872012000200012
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-74. doi: 10.2307/2529310
- Miura, C., Gallani, M., Domingues, G., Rodrigues, R., & Stoller, J. (2010). Adaptação cultural e análise da confiabilidade do instrumento Modified Dyspnea Index para a cultura brasileira. *Revista Latino-Americana de Enfermagem*, 18(5), 1020-1030. doi: 10.1590/S0104-11692010000500025.
- Nakano, T. C. (2013). Problemas apresentados pelos instrumentos com parecer desfavorável no SATEPSI. *Avaliação Psicológica*, 12(2), 121-131. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712013000200003
- Nakano, T. C., & Primi, R. (2012). A estrutura fatorial do Teste de Criatividade Figural Infantil. *Psicologia Teoria e Pesquisa*, 28(3), 275-283. doi: 10.1590/S0102-37722012000300003
- Olea, J., Abad, J. F., & Barrada, J. R. (2010). Tests informatizados y otros nuevos tipos de tests. *Papeles del Psicólogo*, 31(1), 94-107. Recuperado de <http://www.papelesdelpsicologo.es/pdf/1800.pdf>
- Olea, J., Abad, J. F., Ponsoda, V., & Ximenez, C. M. (2004). Un test adaptativo informatizado para evaluar el conocimiento de inglés escrito: Diseño y comprobaciones psicométricas. *Psicothema*, 16(3), 519-525. Recuperado de <https://www.redalyc.org/articulo.oa?id=72716327>
- Olea, J., Ponsoda, V., & Prieto, G. (2006). Revision de libros: Tests informatizados, fundamentos y aplicaciones. *Psicothema*, 12(2), 321-322. Recuperado de <http://www.psicothema.com/psicothema.asp?id=298>
- Pacico, J. C. (2015). Como é feito um teste? Produção de itens. Em C. S. Hutz, D. R. Bandeira, & C. M. Trentini (Eds.), *Psicometria* (pp.55-70). Porto Alegre: Artmed.
- Pasquali, L. (2010). *Instrumentação Psicológica: Fundamentos e práticas*. Porto Alegre: Artes Médicas.
- Perroca, M. G., & Gaidzinski, R. R. (2003). Avaliando a confiabilidade interavaliadores de um instrumento para classificação de pacientes: Coeficiente Kappa. *Revista Escola de Enfermagem USP*, 37(1), 72-80. doi: 10.1590/S0080-62342003000100009
- Prado, O. Z. (2005). Softwares para psicologia: Regulamentação, produção nacional e pesquisas em psicologia clínica. *Boletim de Psicologia*, 55(123), 1-8. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0006-59432005000200006
- Runco, M. A., & Mraz, W. (1992). Scoring divergent thinking tests using total ideational output and a creativity index. *Educational and Psychological Measurement*, 52, 213-221. doi: 10.1177/001316449205200126

Silva, T. F. (2018). Escala Informatizada de Avaliação das Características Criativas: Construção de instrumento (Tese de doutorado, PUC – Campinas). Recuperado de <http://tede.bibliotecadigital.puc-campinas.edu.br:8080/jspui/handle/tede/1074>

recebido em novembro de 2017
aprovado em abril de 2019

Sobre as autoras

Tatiana de Cassia Nakano é psicóloga, doutora em Psicologia pela Pontifícia Universidade Católica de Campinas. Atualmente é docente no Programa de Pós-Graduação Stricto Sensu em Psicologia da Puc-Campinas.

Talita Fernanda da Silva é psicóloga, doutora em Psicologia pela Pontifícia Universidade Católica de Campinas. Atualmente atua como "Consultora Acadêmica", com responsabilidade de Cursos de Graduação relacionados a área de Humanas de uma empresa na área de educação.