

Métodos Atuais para Avaliação da Criatividade: Vantagens e Questionamentos

Tatiana de Cassia Nakano¹

Pontifícia Universidade Católica de Campinas, Campinas-SP, Brasil

RESUMO

Dada a importância da criatividade em diferentes contextos, um grande número de medidas são encontradas, historicamente, na literatura científica. Mais recentemente, os avanços na psicometria e nos modelos teóricos têm permitido o desenvolvimento de novos métodos para avaliação da criatividade, os quais serão enfocados no presente texto. Dentre eles serão apresentados a avaliação subjetiva, *uniqueness score*, Top 2, *average score*, avaliação consensual e *snapshot*, empregados mais recentemente na literatura científica internacional e ainda pouco explorados no Brasil. Exemplo de utilização e críticas a cada um dos métodos, quando existentes, também serão apresentadas.

Palavras-chave: medidas; testes psicológicos; avaliação psicológica.

ABSTRACT – Current methods for evaluating creativity: advantages and questions

Given the importance of creativity in different contexts, a large number of measures can be found in the scientific literature. More recently, advances in psychometry and theoretical models have allowed the development of new methods for evaluating creativity, this being the focus of this text. Among these methods, the subjective evaluation, uniqueness score, Top 2, average score, consensual evaluation and snapshot are presented, with these methods being the most recently used in the international scientific literature and still little explored in Brazil. Examples of usage and critiques of each of the methods, if any, are also presented.

Keywords: measurement; psychological tests; psychological assessment.

RESUMEN – Métodos actuales para la evaluación de la creatividad: ventajas y cuestionamientos

Dada la importancia de la creatividad en diferentes contextos, se encuentran un gran número de medidas en la literatura científica a lo largo de la historia. Por otro lado, los avances más recientes en la psicometría y en los modelos teóricos han permitido el desarrollo de nuevos métodos para la evaluación de la creatividad, los cuales serán investigados en el presente estudio. Entre ellos, se presentará la evaluación subjetiva, *uniqueness score*, Top 2, *average score*, evaluación consensual y *snapshot*, últimamente empleados en la literatura científica internacional y aún poco explorados en Brasil. Ejemplo de utilización y críticas a cada uno de los métodos, cuando existentes, también serán presentados.

Palabras clave: medidas; tests psicológicos; evaluación psicológica.

A criatividade é compreendida como um construto multidimensional, determinado por quatro componentes fundamentais: pessoa, processo, produto e ambiente criativo (Isaksen, Firestien, Murdock, Puccio, & Treffinger, 1994; Rhodes, 1961). Sua avaliação teve início no século XX, a partir do interesse de pesquisadores em obter respostas para uma série de questionamentos: Em quais domínios ou áreas o indivíduo apresenta seu maior potencial? Qual o nível de criatividade que ele apresenta? Quais predições podem ser feitas a partir dos resultados da avaliação? Quais seus pontos fortes e fracos? Como estimular a criatividade de um indivíduo com esse perfil? (Treffinger, Schoonover, & Selby, 2013). Desde então, uma grande diversidade de métodos, instrumentos e técnicas para avaliação da criatividade tem sido proposta, envolvendo formas

tradicionais e novas tendências (Nakano, 2018b). É nesse segundo ponto que o texto se embasa.

Dada a amplitude de possibilidades, diversas tentativas de agrupar as medidas existentes são encontradas na literatura científica (Eysenck, 1999; El-Murad & West, 2004; Hocevar & Bachelor, 1989), as quais envolvem: testes de pensamento divergente, inventários de personalidade, inventários de interesses e atitudes, classificações e indicações feitas por professores, pares e supervisores, julgamento do produto criado, estudos de biografias, inventários biográficos, atividades e realizações criativas relatadas pela própria pessoa. Dentre esses, os testes de pensamento divergente têm sido mais frequentemente utilizados (Benedeck, Muhlmann, Jauk, & Neubauer, 2013).

¹ Endereço para correspondência: Programa de Pós-Graduação em Psicologia. Avenida John Boyd Dunlop, s/n, Jardim Ipaussurama, 13060-904, Campinas, SP. Tel.: (19) 3365-5164. E-mail: tatiananakano@hotmail.com

Historicamente, os modelos mais difundidos e que embasaram vários dos instrumentos disponíveis atualmente foi o desenvolvido por Guilford (1956) e aperfeiçoado por Torrance (1966). O primeiro autor propôs uma série de características cognitivas que estariam presentes nas pessoas criativas sendo que, dentre essas, quatro acabariam por assumir papel central (fluência – número de ideias; flexibilidade = diferentes categorias de ideias; elaboração = adição de detalhes à resposta; originalidade = respostas não comuns) e tidas como base para uma série de instrumentos desenvolvidos. Posteriormente, Torrance (1966) ampliou esse modelo, a partir da adição de características emocionais (expressão de emoção, fantasia, movimento, dentre outras). A partir de seu modelo, as características descritoras da pessoa criativa passaram a ser divididas em características cognitivas e emocionais, sendo importante destacar a relevância de ambas para a expressão criativa. Tais modelos serviram de base para a construção de diferentes instrumentos de avaliação da criatividade, os quais, usualmente se baseiam na soma da pontuação obtida pelo sujeito em cada uma dessas características, em diferentes tarefas.

De forma mais frequente, a avaliação do construto envolve a apresentação de uma tarefa não usual, sendo solicitado que o indivíduo forneça o maior número de ideias possíveis (Chu & Lin, 2013), em um contexto específico, dentro de um período de tempo limitado e padronizado de resposta. Dentre os instrumentos disponíveis, uma série de elementos relacionados à criatividade vem sendo incluídos, tais como pensamento divergente, pensamento analítico, flexibilidade mental, pensamento associativo, combinação seletiva, aspectos relacionados à personalidade (tolerância à ambiguidade, tomada de risco, abertura a experiências), pensamento intuitivo, motivação para criar (Lubart, Zenasni, & Barbot, 2013), ou ainda espontaneidade, imaginação e capacidade inventiva (Nogueira & Bahia, 2006).

Tal avaliação pode ser realizada, segundo Nakano (2018), a partir de diferentes métodos: informal (aquele que pode ser realizado a qualquer tempo e em qualquer lugar, a partir da observação dos comportamentos do sujeito), formal (envolvendo a coleta de informações por meio de instrumentos padronizados, aplicados de forma coletiva ou individual), qualitativos (englobam observações, dados biográficos, entrevistas, análise de produção criativa, visando a identificação de indicadores de criatividade e das circunstâncias que favorecem ou dificultam a expressão criativa) e quantitativos (embasam-se na utilização de instrumentos padronizados, cujos resultados permitem que se classifique o nível de habilidade de um sujeito, por meio de um valor numérico, geralmente média e percentil). Usualmente, como forma de avaliar a extensão com que a pessoa produziu um trabalho considerado criativo naquele contexto, seu desempenho é comparado com aquele obtido por outros indivíduos, a fim de que seu potencial possa ser estimado (Lubart et al., 2013).

Também podem ser classificados em métodos objetivos e subjetivos. Os primeiros envolvem uma série de indicadores, tais como aqueles propostos por Guilford (1956) e Torrance (1966): fluência, flexibilidade, etc. Em relação aos subjetivos, a autoavaliação, julgamento de produtos por juízes, indicações realizadas por pares, avaliação de produtos tem sido mais utilizados. A literatura tem recomendado a combinação entre as técnicas objetivas e subjetivas como a melhor opção para a avaliação dos testes de criatividade (Plucker, Qian, & Wang, 2011).

Entretanto, indiferente ao método selecionado, deve-se ter clareza de que, ao se avaliar um construto tão complexo como a criatividade, nenhuma medida, isoladamente, conseguirá cobrir todas as suas dimensões, de modo que os resultados devem ser restritos ao tipo de criatividade que o instrumento utilizado se propõe a avaliar (Nakano, 2018). Consequentemente, pesquisadores têm debatido qual seria a forma mais adequada de avaliação, não tendo sido alcançado, até o momento, um consenso (Beghetto, Plucker, & MaKinster, 2001; Cropley, 2000).

Isso porque, dentre os desafios a serem enfrentados pela área, alguns se mostram essenciais: 1. desfazer a ideia de que a criatividade pode ser tomada como um domínio único e seu resultado como indicador de desempenho geral, passível de ser avaliado por meio de um único instrumento e das tarefas nele envolvidas (Kim, 2011), 2. ampliar a compreensão acerca da existência de uma grande diversidade de domínios específicos e suas diferentes manifestações (por exemplo, na área verbal, figural, corporal, musical e que podem representar diferentes facetas da criatividade), 3. necessidade de rever o sistema de pontuação dos instrumentos existentes, visto que, na maior parte deles, ainda que diversos tipos de tarefas e características sejam avaliadas, as notas finais são posteriormente combinadas por meio de sua soma ou média simples, como se fosse possível se chegar a uma pontuação única e final sobre a criatividade do indivíduo (Silvia, 2011). Segundo esse autor, tais questões desconsideram os diferentes processos criativos envolvidos em cada tipo de tarefa, de modo a afirmar sua equivalência e ignorar a possibilidade de que cada uma capture diferentes facetas da criatividade, assim como conhecimentos, estratégias e habilidades diferenciadas.

Devido à essas controvérsias e à tentativa dos pesquisadores de conhecer os limites das medidas frente a um construto tão amplo, discussões recentes acerca dos métodos utilizados na sua avaliação se tem feito presentes, considerando-se, principalmente, questões relacionadas aos métodos utilizados na pontuação das medidas de criatividade (Nakano, 2018a). Do mesmo modo, o fato de que, indiferente aos avanços na psicometria e na avaliação nos últimos 40 anos, as antigas formas de avaliação ainda são utilizadas (Silvia, 2008), apesar de algumas melhorias e alternativas desenvolvidas desde então. Tais avanços tornam injustificável a utilização

das formas tradicionais, sem questionamento ou revisão delas.

Consequentemente, novos métodos para avaliação da criatividade têm sido encontrados na literatura científica, perante a necessidade de desenvolvimento de novos modelos de avaliação, mais otimizados, válidos e precisos. Tais propostas serão apresentadas a seguir, sendo focado o método intitulado avaliação subjetiva, as técnicas *uniqueness score*, *Top 2*, *average score*, avaliação consensual e *snapshot* (usualmente utilizados na literatura internacional e ainda pouco difundidos no Brasil).

Tais modelos propõem, tanto a utilização de novas formas de avaliação, como de métodos mais recentes de análise dos dados, tais como modelagem multinível (Silvia, 2007), emprego de modelos de créditos parciais (Nakano & Primi, 2014) ou do *Many-Facet Rasch* (Primi, 2014). Exemplo de utilização de cada método será apresentado, sendo importante salientar que, devido ao fato de ainda se mostrarem recentes, as pesquisas que fizeram uso deles limitam-se, em sua maior parte, à investigação de suas evidências de validade ou precisão na avaliação da criatividade. Também serão apresentadas, quando existentes, críticas a eles, de forma a possibilitar um julgamento mais amplo por parte do leitor.

Uniqueness Score

Nesse método, juízes externos avaliam todas as respostas dadas, pelo sujeito, em relação à sua originalidade e adequabilidade, sendo as pontuações somadas para dar origem a uma pontuação total. Proposto inicialmente por Wallach e Kogan (1965), prevê a pontuação das respostas dos sujeitos a testes de pensamento divergente em dois escores: fluência (número de respostas) e a singularidade (número de respostas únicas, ou seja, originalidade). Assim receberia um ponto aquelas respostas que não tivessem sido dadas por nenhum outro participante, sendo, as demais, pontuadas em zero. Dada sua simplicidade perante os demais métodos de correção existentes (os quais envolvem um número muito maior de características criativas avaliadas), o método da singularidade tornou-se popular dada a facilidade na sua cotação e sua objetividade (Silvia, 2015).

Como exemplo de utilização desse método, pode-se citar o estudo conduzido por Cropley, Kaufman e Cropley (2011). No estudo, o método do *Uniqueness Score* foi utilizado para verificar as evidências de validade e precisão de um instrumento denominado *The Creative Solution Diagnosis Scale* (CSDS). A amostra foi composta por 323 estudantes universitários, os quais foram apresentados a imagens de cinco ratoeiras. Os participantes foram convidados a classificar cada uma delas de acordo com os itens da escala CSDS, dentro de uma escala Likert de 5 pontos (*nada criativo a muito criativo*), para indicar o grau de criatividade de cada ratoeira.

A precisão de juízes foi avaliada por meio do alfa de Cronbach, o qual indicou um valor de 0,948 para o total

do instrumento e valores oscilando entre 0,851 e 0,961 para cada estímulo apresentado. A análise fatorial confirmatória foi utilizada para verificar a estrutura fatorial, se ela se relacionava a um modelo unidimensional (no caso a criatividade) ou se a uma série de outras variáveis, tais como medidas ou itens individuais). Os resultados indicaram a existência de cinco fatores: relevância e efetividade, problematização/novidade, propulsão, elegância e gênese. Segundo os autores, os dados apontaram para a possibilidade de que avaliadores não especializados possam ser capazes de alcançar um entendimento comum do que significa criatividade, reconhecer tal característica quando ela é expressa e concordar uns com os outros sobre essa classificação, de modo que, no estudo apresentado, o *Uniqueness Score* se mostrou um método útil, válido e preciso na investigação das qualidades psicométricas de um instrumento em processo de desenvolvimento.

As críticas ao modelo amparam-se em três pontos principais (Silvia et al., 2008). O primeiro refere-se à constatação da dependência da amostra, visto que, dentro de uma amostra pequena a probabilidade de resposta única é muito maior do que aquela encontrada em grandes amostras. Desse modo, a prevalência de respostas únicas (singulares) declinaria de acordo com o aumento do tamanho da amostra, variando entre os estudos, de maneira a dificultar a generalização dos resultados e sua comparabilidade. O segundo ponto marca-se pela possível associação entre fluência e singularidade visto que pessoas que geram mais ideias, provavelmente apresentam maior probabilidade de ter respostas únicas. Baseando-se nesse argumento, as duas pontuações não seriam independentes. O terceiro ponto, por sua vez, ampara-se no fato de que tal método avalia a quantidade de ideias, não sua qualidade. Consequentemente, parte das respostas que poderiam ser pontuadas em singularidade e, portanto, consideradas criativas, não necessariamente corresponderiam a respostas com alta qualidade, podendo se caracterizar como irrelevantes ou inapropriadas para a tarefa. Ao não separar as respostas únicas daquelas que são inadequadas, aumenta-se o erro da medida e, como consequência, a possibilidade de falsos positivos para criatividade.

Top 2 e Average Score

Partindo-se da combinação de algumas propostas já existentes, alguns novos métodos subjetivos de avaliação da criatividade têm sido testados. Diferentemente dos métodos tradicionais que envolvem a cotação de diversas características associadas à criatividade, anteriormente citadas, os dois novos métodos consistiriam, segundo Silvia (2008) em: 1. *Average Score*, ou escore médio, no qual todas as respostas dadas pelo sujeito são estimadas, calculando-se uma média simples das pontuações e 2. *Top 2*, o qual envolveria a pontuação das duas melhores respostas em cada tarefa, indicadas pelo próprio examinando. Em ambos os métodos é utilizada uma escala de 1 a 5 pontos (1=nada criativo, 5=muito criativo).

As respostas do sujeito, nesse modelo, seriam avaliadas a partir do julgamento de critérios como raridade, distância entre as associações e inteligência (julgada por meio do uso de recursos como humor, ironia, combinação e perspicácia).

Ao selecionar suas melhores respostas, Benedeck et al. (2013) apontam, como principais vantagens desse método, o fato de que as ideias escolhidas pelo próprio sujeito irão refletir a máxima performance do sujeito, diminuir o tempo e recurso gasto pelos juízes com a avaliação do número total de ideias em cerca de 70%, causar menos fadiga nos sujeitos e propiciar escores mais acurados, principalmente em relação à originalidade. Isso porque, segundo Plucker et al. (2011), o uso de menos ou muito mais ideias pode resultar em escores menos válidos. Opções alternativas, como uso da melhor ideia (*Top 1*) ou das três melhores ideias (*Top 3*) também vêm sendo testadas (Benedeck, Franz, Heene, & Neubauer, 2012), sendo importante destacar que, segundo os autores, o número mais adequado de ideias também irá depender do tempo da tarefa.

Como exemplo de utilização desses métodos, Silvia (2008) conduziu uma pesquisa com 226 estudantes universitários dos cursos de Psicologia, Enfermagem e Biologia, que faziam parte de um projeto de criatividade e cognição. Os participantes responderam a tarefas de pensamento divergente: duas de usos incomuns (para um tijolo e uma faca), duas tarefas de instância (coisas que são redondas e que fazem barulho) e duas tarefas de consequências (de não precisar dormir e se todos fossem encolhidos). Nas instruções, as pessoas eram informadas de que as tarefas envolviam criatividade e que elas deveriam tentar ser criativas. Após o término da tarefa, limitada a três minutos cada, os participantes eram solicitados a lerem suas respostas e circularem as duas que achavam que eram as melhores.

Três estudantes de graduação atuaram como juízes, avaliando, independentemente, cada resposta selecionada pelo próprio participante, sendo importante salientar que eles não tiveram acesso às respostas que não foram escolhidas. A escala variava entre 1 e 5 (1=nada criativa; 5=muito criativa) no método do *Top 2*. Para o método *Average Score*, cada resposta recebeu uma classificação e a média da pontuação foi calculada, sendo considerada como a pontuação final do sujeito. Os resultados indicaram que o método do *Average Score* se mostrou mais preciso do que o *Top 2*.

Esses métodos foram recebidos com bastante crítica no meio científico, sendo alvo de questionamentos relacionados à metodologia utilizada (Kim, 2008; Lee, 2008; Mumford, Vessey, & Barrett, 2008), bem como às suas interpretações sobre o conceito de criatividade envolvido (Runco, 2008) e uso inadequado da técnica de avaliação consensual (Baer, 2008), devendo-se salientar que opiniões favoráveis, ainda que em número bem menor, também foram encontradas (Kogan, 2008).

Os questionamentos envolvem dúvidas em relação à acurácia ou precisão desse julgamento, visto que uma série de variáveis poderia exercer influência no processo de correção, envolvendo, principalmente, a questão da subjetividade (uma vez que envolve o julgamento das respostas por juízes, assim como a avaliação de aspectos sujeitos a interpretações pessoais, tais como humor e perspicácia).

Do mesmo modo, as críticas envolvem a discrepância no julgamento entre avaliadores, assim como a ausência de preparo do próprio respondente para selecionar suas duas melhores respostas, no caso do *Top 2* (Silvia, 2008). Sobre tal argumento, Ng e Feldman (2012) sugerem que outros estudos mostraram que o próprio indivíduo se encontra na melhor posição, do que dos próprios juízes, de julgar sua própria criatividade, visto que ele é a pessoa mais consciente do valor de suas contribuições.

Em resposta às críticas, Silvia, Winterstein e Willse (2008) reafirmam a importância desse tipo de estudo, ressaltando que investigações futuras poderão revelar os pontos fortes e fracos dessa abordagem, justificando sua relevância ao considerarem que a Psicologia da criatividade tem, historicamente, bastante conhecimento acumulado acerca de como as pessoas geram ideias criativas, mas muito a ser explorado em relação a como elas julgam suas próprias ideias (Silvia, 2008). Entretanto, Benedeck et al. (2013) chamam a atenção para a influência da quantidade de tempo fornecida para realização da atividade sobre o número de ideias fornecidas pelo sujeito e, consequentemente, nas medidas de fluência e originalidade.

Outra questão importante de ser ressaltada refere-se ao fato de que tais métodos têm recebido menos atenção dos pesquisadores se comparados aos métodos objetivos, visto que exigem mais recursos de pesquisa (dada a necessidade de se estimar a não só a variância devida aos traços e níveis de criatividade entre os examinados mas também aquela proveniente da diferença no julgamento entre juízes) e devido ao fato dos pesquisadores ainda confundirem o termo “objetivo” com “válido” (Silvia, 2011). Nesse caso, o termo “métodos subjetivos” já acaba por gerar um estereótipo de ausência de cientificidade, o que não é verdade.

Consensual Assessment Technique (CAT)

Outra técnica bastante utilizada atualmente é a *Consensual Assessment Technique (CAT)*, a qual consiste no julgamento independente de produtos, por juízes especialistas (profissionais em seus domínios específicos), dentro de uma escala que varia entre o mais baixo e o mais alto nível de criatividade. Parte-se da percepção individual de cada juiz acerca do que é criatividade, sem que maiores orientações, definições ou treinamento sejam fornecidos a eles. Nesse método, juízes avaliam cada resposta do sujeito dentro de uma escala, usualmente de 2 a 3 pontos. Essas avaliações usualmente são somadas e dão origem a uma pontuação média.

A pioneira na utilização dessa técnica foi Amabile (1982), baseando-se na definição de que um produto ou resposta mostra-se criativa na medida em que observadores independentes e apropriados (familiarizados com o domínio no qual o produto foi criado) concordam que é criativo. A partir da condução de uma série de estudos que fizeram uso de uma grande diversidade de tarefas (escrita de poemas e histórias, escrita de histórias a partir de figuras, criação de colagens, criação de problemas matemáticos e quebra-cabeças), tendo crianças e adultos como sujeitos, os resultados obtidos pela pesquisadora apontaram para coeficientes de precisão em torno de 0,70 a 0,90 entre diferentes avaliadores, de forma a estimular novos estudos baseados nessa técnica.

A técnica da avaliação consensual tem se mostrado uma ferramenta importante para avaliar a criatividade, sendo amplamente utilizada no contexto internacional (Baer & McKool, 2009). Segundo os autores, ao contrário da maioria dos testes de criatividade, essa técnica não mede habilidades ou traços que são supostamente parte do pensamento criativo ou o potencial do indivíduo. Ela avalia o desempenho criativo real, nos mais diferentes níveis, desde a criatividade do dia a dia quanto nos seus mais altos níveis. De acordo com Baer e McKool (2014), a CAT é, atualmente, a ferramenta que mais tem apresentado evidências de validade para avaliar a criatividade, sendo chamada, por esse motivo, de padrão ouro.

O critério que embasa o CAT se mostra claro e simples, sendo essa a principal vantagem desse método, apesar da exigência de especialistas (Chu & Lin, 2013). No entanto, segundo os autores, outros estudos vêm sendo conduzidos com o objetivo de melhorar esse sistema de pontuação usando, por exemplo, estudantes universitários como juízes.

Como vantagem, seus defensores argumentam que, diferentemente das outras medidas de criatividade, a CAT não seria baseada em uma teoria específica de criatividade, de maneira que sua validade não dependeria desse fator (Kaufman et al., 2010), permitiria a avaliação de uma vasta gama de produtos criativos, sendo sensível a variações culturais e históricas no julgamento de criatividade (Chen et al., 2002), está livre de preconceitos de gênero, raça e etnia (Baer & McKool, 2009), assim como não faz uso de escores padronizados, fato considerado como vantagem dado o fato de que o método realiza a comparação de desempenho entre os participantes, não com uma amostra normativa (Kaufman, Plucker, & Russell, 2012).

Diferentes pesquisas foram conduzidas a fim de investigar aspectos relacionados à capacitação dos juízes e a necessidade de serem ou não especialistas na área de julgamento. Os resultados mostraram diferenças importantes na avaliação feita por juízes especialistas e não especialistas no julgamento de produtos, confirmada por meio de baixa correlação encontrada entre os grupos (Kaufman, Baer, Cole, & Sexton, 2008), bem como no

fato de que os julgamentos feitos pelos *experts* se mostraram mais consistentes do que aqueles realizados por novatos na área (Kaufman, Baer, & Cole, 2011). Tais resultados parecem indicar que a substituição do juiz especialista na área não se mostra uma tarefa simples, correndo-se o risco de que tal procedimento possa derivar em diferentes resultados, sendo importante considerar, ainda de acordo com os autores, o fato de que juízes *experts* apresentam, de uma forma geral, maior rigidez e padrões mais elevados na avaliação dos produtos, usualmente fornecendo notas mais baixas que os não *experts* (Kaufman, Lee, Baer, & Lee, 2007).

Como exemplo de utilização desse método, pode-se citar o estudo conduzido por Chu e Lin (2013). Na pesquisa, trinta e sete alunos do ensino fundamental foram convidados a completarem um teste de pensamento divergente e uma tarefa de escrita criativa. Ambas as atividades foram corrigidas por dois juízes especialistas, os quais deveriam julgar as respostas em relação à criatividade, por meio da técnica de avaliação consensual. A correlação entre as avaliações foi de 0,73 para a criatividade. A correlação entre o julgamento dos índices tradicionais do teste de pensamento divergente com aquele obtido via CAT apontou valores significativos para fluência ($r=0,53$), originalidade ($r=0,59$), flexibilidade ($r=0,37$) e a pontuação total ($r=0,56$).

As críticas ao modelo envolvem a dificuldade de se conseguir, em número suficiente, especialistas no domínio a ser investigado quando uma grande quantidade de produtos precisa ser avaliada, além do seu alto custo, ausência de estudos investigando a eficácia do treinamento de juízes não *experts* a fim de que investigar se, dessa maneira, eles conseguem realizar julgamentos similares aos dos *experts* (Kaufman et al., 2008), além das dificuldades causadas pela utilização de um pressuposto de que há apenas uma faceta na avaliação da criatividade, no caso a variância entre juízes, desconsiderando-se a influência de outros fatores na determinação da precisão da medida (Kaufman et al., 2007).

Outras críticas ao modelo, destacadas por Piffer (2012), amparam-se no grau moderado de confiabilidade teste-reteste (que não é muito elevado em comparação com a maioria dos instrumentos psicométricos), um alto grau de especificidade de domínio (baixas correlações entre o desempenho criativo em diferentes domínios), bem como à avaliação de um número restrito de produtos. Desse modo, argumenta que a CAT pode avaliar o potencial criativo em domínios específicos e não como um traço geral. Também Amabile (1982) aponta para o fato de que esse método mostra-se impraticável em curto prazo; sua restrição ao momento histórico e às concepções de criatividade dominantes, destacando ainda a dificuldade de sua aplicação no julgamento de produtos revolucionários, dada a dificuldade dos juízes em chegar a um acordo sobre o nível de criatividade evidente dada a ausência de familiaridade com o novo domínio. Também

Cseh e Jeffries (2019) apontam que o profundo conhecimento sobre os procedimentos dessa técnica, bem como seus efeitos, ainda permanece incompleto, de modo que lacunas ainda se fazem presentes.

Snapshot

Outro método recente na literatura é o *Snapshot* (Tiro Rápido), pesquisado por Silvia, Martin e Nusbaum (2009), considerado um método rápido e simples para avaliar a criatividade e que, por esses motivos, visa substituir os demais, dadas as vantagens relacionadas à economia de tempo e pessoal necessário. Nessa proposta, os avaliadores julgam um conjunto de respostas e fornecem uma classificação global única para todo o conjunto. Nenhuma codificação interna ou pontuação é executada, evitando-se a necessidade de transcrição ou digitação das respostas para uma folha de correção/pontuação, de modo a facilitar o trabalho do avaliador.

Como exemplo de utilização desse método, pode-se citar o estudo conduzido por Silvia et al. (2009). Estudantes de Psicologia ($n=226$) responderam a uma série de tarefas de pensamento divergente e escalas de autorrelato. Três juízes avaliaram as 1356 respostas dadas pelos sujeitos, classificando o conjunto total de respostas de cada sujeito em uma escala de 1 (*nada criativo*) a 5 (*muito criativo*). Os resultados indicaram que a precisão entre avaliadores apresentou valor de 0,83, bem como evidências de validade concorrente foram confirmadas por meio da capacidade desse método em prever o escore dos sujeitos na medida de abertura à experiência.

Considerando-se que o método ainda não se encontra bem explorado, nenhuma crítica foi encontrada na literatura científica. Entretanto, os próprios autores defendem que pesquisas futuras poderão investigar, de forma mais aprofundada, a relação entre conveniência e efetividade psicométrica. Caso confirmadas, argumentam que tal método apresentará diversas vantagens frente aos demais, os quais envolvem mais esforço, ressaltando, dentre elas, a possibilidade de utilização de um método também válido e preciso, mais simples e rápido.

A Utilização da Tecnologia na Avaliação da Criatividade

Acompanhando as tendências de utilização dos recursos da informatização e das tecnologias de informação e comunicação (TICs) na avaliação psicológica, algumas propostas têm sido testadas na avaliação da criatividade. Como exemplo pode ser citado o processo de treinamento de juízes para correção de uma nova ferramenta: *Evaluation of Potential Creativity* (EPOC), voltada para avaliação do potencial criativo de crianças e adolescentes (Lubart, Besançon, & Barbot, 2011). O instrumento, mais recentemente, teve um sistema *on-line* de correção e pontuação, bem como de treinamento de juízes, testado.

O sistema é apresentado em versões em inglês, francês, alemão, turco e árabe, sendo, as normas,

constantemente revisadas e atualizadas, dentro de um processo dinâmico, conforme novas respostas são obtidas. O sistema também oferece uma nova perspectiva no treinamento de juízes na avaliação das respostas. Por meio de um *website* interativo, os juízes podem visualizar respostas de sujeitos e sugerirem uma pontuação para elas. Automaticamente o sistema oferece um *feedback* para esse juiz. Dessa maneira, os juízes podem aprender de forma interativa e desenvolver uma diversidade de conhecimentos acerca do tipo de resposta que as crianças podem apresentar. Depois de uma sequência de treinamento, o juiz pode testar sua habilidade por meio da avaliação de um conjunto de respostas, previamente corrigidas por juízes mais experientes (Lubart et al., 2013).

Outro exemplo é relatado por Beketayev e Runco (2016). Os pesquisadores desenvolveram um método algorítmico automatizado baseado na análise semântica para avaliar o pensamento divergente. Nele, os examinados recebem as tarefas em um computador ou dispositivo móvel e suas respostas são automaticamente comparadas com normas e redes semânticas. Ao compararem a pontuação de 250 sujeitos em flexibilidade estimada por meio desse método, com a pontuação estimada por meio dos métodos tradicionais, os resultados confirmaram uma correlação de 0,74 entre elas de modo, que segundo os autores, tais resultados podem justificar a escolha pelo método informatizado quando se trata da investigação em grande escala, dado o tempo de computação quase imediato dos resultados.

Apesar dos exemplos internacionais citados, convém destacar que as TICs ainda vêm sendo utilizadas, no Brasil, na temática da criatividade muito mais relacionada ao seu uso como facilitador da expressão criativa na educação básica (Borges & Fleith, 2018; Vilarinho-Rezende, Borges, Fleith, & Joly, 2016) do que na área da avaliação psicológica.

Considerações Finais

Destacada como uma das *21st Century Skills*, juntamente com o pensamento crítico, colaboração e comunicação (Kaufman, 2010; Piirto, 2011; Treffinger et al., 2013), a criatividade vem sendo, cada vez mais, valorizada, ainda que a revisão da literatura aponte para várias questões históricas que vêm sendo enfrentadas pelos pesquisadores interessados na sua avaliação. Dificuldades na sua identificação ainda se fazem presentes, dada a ausência de uma definição consensual, assim como a grande diversidade de métodos que vêm sendo utilizados pelos pesquisadores, das mais diversas áreas, na tentativa de definir a forma mais adequada de avaliação desse construto.

Como resultado, diversos métodos têm sido propostos por pesquisadores, sendo que, no entanto, embora com resultados promissores, as pesquisas ainda se mostram iniciais. Nenhum dos métodos apresentados conta,

até o momento, com número suficiente de pesquisas que comprovem, de forma satisfatória, sua eficácia ou evidências de validade. Conseqüentemente, suas investigações necessitam de aprofundamento e continuidade. Importante salientar que os métodos aqui abordados se encontram em uso no contexto internacional não tendo sido encontrada, até o momento, nenhum estudo brasileiro que tenha feito uso deles.

Não se pode deixar de citar que parte dessas inovações é resultado dos avanços nos métodos estatísticos empregados nas análises psicométricas, de maneira que o uso desses recursos tem possibilitado, aos pesquisadores, apontar direções futuras para pesquisa e prática na área (Plucker & Runco, 1998).

Entretanto, estudos na temática ainda mostram-se necessários, principalmente com a finalidade de contemplar áreas da criatividade ainda escassas (podendo-se citar, por exemplo, criatividade musical ou corporal), outros conceitos relativamente novos associados à criatividade (criatividade emocional, inteligência criativa e liderança criativa), assim como aqueles voltados à investigação da aplicação contextualizada da criatividade aos diferentes propósitos intelectuais, sociais e práticos (Silvia, 2014). Somente por meio da ampliação das pesquisas, tais propostas poderão reduzir as desconfiças dos pesquisadores, de modo a constituírem-se em métodos reconhecidamente inovadores e aceitos pela comunidade científica.

Referências

- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43, 997-1013. doi: 10.1037/0022-3514.43.5.997
- Baer, J. (2008). Divergent thinking tests have problems, but this is not the solution. *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 89-92. doi: 10.1037/1931-3896.2.2.89
- Baer, J., & McKool, S.S. (2009). Assessing creativity using the Consensual Assessment Technique. Em C. Schrieder (Eds.), *Handbook of research on assessment technologies, methods and applications in higher education* (pp. 65-77). New York: Information Science Reference.
- Baer, J., & McKool, S.S. (2014). The gold standard for assessing creativity. *International Journal of Quality Assurance in Engineering and Technology Education*, 3(1), 81-93. doi: 10.4018/ijqaete.2014010104
- Beghetto, R. A., Plucker, J. A., & MaKinster, J. G. (2001). Who studies creativity and how do we know? *Creativity Research Journal*, 13(3/4), 351-357. doi: 10.1207/S15326934CRJ1334_12
- Beketayev, K., & Runco, M. A. (2016). Scoring divergent thinking tests by computer with a semantics-based algorithm. *European Journal of Psychology*, 12(2), 201-220. doi: 10.5964/ejop.v12i2.1127
- Benedeck M., Franz F., Heene M., & Neubauer A. C. (2012). Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences*, 53, 480-485.
- Benedeck, M., Mulhmann, C., Jauk, E., & Neubauer, A.C. (2013). Assessment of divergent thinking by means of the subjective top-scoring method: Effects of the number of top-ideas and time-on-task on reliability and validity. *Psychology of Aesthetics, Creativity and the Arts*, 7(4), 341-349. doi: 10.1037/a0033644
- Borges, C. N., & Fleith, D. S. (2018). Uso da tecnologia na prática pedagógica: Influência na criatividade e motivação de alunos do ensino fundamental. *Psicologia: Teoria e Pesquisa*, 34, e3435. doi: 10.1590/0102.3772e3435
- Chen, C., Kasof, J., Himsel, A.J., Greenberger, E., Dong, Q., & Xue, G. (2002). Creativity in drawings of geometric shapes: A cross-cultural examination with the consensual assessment technique. *Journal of Cross-Cultural Psychology*, 33(2), 171-187. doi: 10.1177/0022022102033002004
- Chu, T., & Lin, W. (2013). Uniqueness, integration or separation? Exploring the nature of creativity through creative writing by elementary school students in Taiwan. *International Journal of Experimental Educational Psychology*, 33(5), 582-595. doi: 10.1080/01443410.2013.821459
- Cropley, A. J. (2000). Defining and measuring creativity: Are creativity tests worth using? *Rouper Review*, 23(2), 72-80. doi: 10.1080/02783190009554069
- Cropley, D. H., Kaufman, J. C., & Cropley, A. J. (2011). Measuring creativity for innovation management. *Journal of Technology Management & Innovation*, 6(3), 13-29. Recuperado de <https://www.redalyc.org/articulo.oa?id=84720911002>
- Cseh, G. M., & Jeffries, K. K. (2019). A scattered CAT: A critical evaluation of the consensual assessment technique for creativity research. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 159-166. doi: 10.1037/aca0000220
- El-Murad, J., & West, D.C. (2004). The definition and measurement of creativity: What do we know? *Journal of Advertising Research*, 44(2), 188-201. doi: 10.1017/S0021849904040097
- Eysenck, H. (1999). As formas de medir a criatividade. Em M.A. Boden (Eds.). *Dimensões da criatividade* (pp. 203-225). Porto Alegre: Artes Médicas.
- Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, 53(4), 267-293.
- Hocevar, D. & Bachelor, P. (1989). A taxonomy and critique of measures used in the study of creativity. In J. A. Glover, R. R. Ronning & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 53-75). New York: Plenum.
- Isaksen, S. G., Firestien, R. L., Murdock, M. C., Puccio, G. J. & Treffinger, D. L. (1994). *The assessment of creativity*. Buffalo, New York: Center for Creative Studies.
- Kaufman, J. C. (2010). Using creativity to reduce ethnic bias in college admissions. *Review of General Psychology*, 14, 189-203. doi:10.1037/a0020133

- Kaufman, J. C., Baer, J., Agars, M. D., & Loomis, D. (2010). Creativity stereotypes and the consensual assessment technique. *Creativity Research Journal*, 22(2), 200-205. doi: 10.1080/10400419.2010.481529
- Kaufman, J. C., Baer, J., & Cole, J. C. (2011). Expertise, domains, and the Consensual Assessment Technique. *The Journal of Creative Behavior*, 43(4), doi: 10.1002/j.2162-6057.2009.tb01316.x
- Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J.D. (2008). Comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal*, 20(2), 171-178. doi: 10.1080/10400410802059929
- Kaufman, J. C., Lee, J., Baer, J., & Lee, S. (2007). Captions, consistency, creativity and the consensual assessment technique: New evidence of reliability. *Thinking Skills and Creativity*, 2, 96-106. doi:10.1016/j.tsc.2007.04.002
- Kaufman, J. C., Plucker, J. A., & Russell, C. M. (2012). Identifying creativity as a component of giftedness. *Journal of Psychoeducational Assessment*, 30, 60-73. doi: 10.1177/0734282911428196
- Kim, K. H. (2008). The Torrance Tests of Creative Thinking already overcome many of the perceived weakness that Silvia et al.'s (2008) methods are intended to correct. *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 97-99. doi: 10.1037/1931-3896.2.2.97
- Kim, K. H. (2011). The APA 2009 division 10 debate: Are the Torrance Tests of Creative Thinking still relevant in the 21st century? *Psychology of Aesthetics, Creativity, and the Arts*, 5(4), 302-308. doi: 10.1037/a0021917
- Kogan, N. (2008). Divergent-thinking research and the Zeitgeist. *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 100-102. doi: 10.1037/1931-3896.2.2.100
- Lee, S. (2008). Reliability and validity of uniqueness scoring in creativity assessment. *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 103-108. doi: 10.1037/1931-3896.2.2.103
- Lubart, T.I., Besançon, M. & Barbot, B. (2011). *EPoC: Evaluation du potentiel créatif* (Test et manuel). Paris: Hogrefe.
- Lubart, T., Zenasni, F., & Barbot, B. (2013). Creative potential and its measurement. *International Journal form Talent Development and Creativity*, 1(2), 41-50. Recuperado de https://www.researchgate.net/profile/Baptiste_Barbot/publication/264618782_Creative_Potential_and_Its_Measurement/links/53e923a90cf2dc24b3c8937b.pdf
- Mumford, M. D., Vessey, W. B., & Barrett, J. D. (2008). Measuring divergent thinking: Is there really one solution to the problem? *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 86-88. doi: 10.1037/1931-3896.2.2.86
- Nakano, T.C. (2018a). Recent considerations on creativity assessment methods. In J.C. Penagos-Corzo & M.A.P. Vargas (Eds.), *Challenges in creativity & psychology for the XXI century* (pp. 31-42). Guadalajara: Universidade de las Américas Puebla & Universidad de Guadalajara.
- Nakano, T. C. (2018b). A criatividade pode ser medida? Reflexões sobre métodos utilizados e questões envolvidas. *Arquivos Brasileiros de Psicologia*, 70(1), 128-145. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-52672018000100010
- Nakano, T. C., & Primi, R. (2014). Rasch-Master's Partial Credit Model in the Assessment of Children's Creativity in Drawings. *The Spanish Journal of Psychology*, 17, 1-16. doi: 10.1017/sjp.2014.36
- Ng, T. W. H., & Feldman, D. C. (2012). A comparison of self-ratings and non self-report measures of employee creativity. *Human Relations*, 65(8), 1021-1047. doi: 10.1177/0018726712446015
- Nogueira, S. I., & Bahia, S. (2006). Avaliação da criatividade ou a necessária criatividade na avaliação. *Revista Lusófona de Ciências da Mente e do Comportamento*, 1(1), 47-88. Recuperado de <http://recil.grupolusofona.pt/bitstream/handle/10437/286/Saralbe.pdf?sequence=1>
- Piffer, D. (2012). Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research. *Thinking Skills and Creativity*, 7, 258-264. doi: 10.1016/j.tsc.2012.04.009
- Piirto, J. (2011). *Creativity for 21st century skills: How to embed creativity into the curriculum*. Netherlands: Sense Publishers.
- Plucker, J. A., Qian, M., & Wang, S. (2011). Is originality in the eye of the beholder? Comparison of scoring techniques in the assessment of divergent thinking. *Journal of Creative Behavior*, 45, 1-22. doi: 10.1002/j.2162-6057.2011.tb01081.x
- Plucker, J. A., & Runco, M. A. (1998). The death of creativity measurement has been greatly exaggerated: Current issues, recent advances, and future directions in creativity assessment. *Roepers Review*, 21(1), 36-40. doi: 10.1080/02783199809553924
- Primi, R. (2014). Divergent productions of metaphors: Combining many-facet rasch measurement and cognitive psychology in the assessment of creativity. *Psychology of Aesthetics, Creativity, and the Arts*. doi: 10.1037/a0038055
- Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappa*, 42, 305- 310.
- Runco, M. A. (2008). Divergent thinking is not synonymous with creativity. *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 93-96. doi: 10.1037/1931-3896.2.2.93
- Silvia, P. J. (2007). An introduction to multilevel modeling for research on the psychology of art and creativity. *Empirical Studies of the Arts*, 25(1), 1-20. doi: 10.2190/6780-361T-3J83-04L1.
- Silvia, P. J. (2008). Discernment and creativity: How well can people identify their most creative ideas? *Psychology of Aesthetics, Creativity, and the Arts*, 2(3), 139-146. doi: 10.1037/1931-3896.2.3.139
- Silvia, P. J. (2011). Subjective scoring of divergent thinking: Examining the reliability of unusual uses, instances and consequences tasks. *Thinking Skills and Creativity*, 6, 24-30. doi: 10.1016/j.tsc.2010.06.001
- Silvia, P. J. (2014). Why Big Theories are fruitless, fragmentation is ideal, defining creativity is overrated and method-driven research is urgent: Some thoughts on the flourishing state of creativity science. *Creativity: theories, research, applications*, 1(2), 233-239. doi: 10.15290/cra.2014.01.02.10
- Silvia, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychological Review*, 27(4), 599-606. doi 10.1007/s10648-015-9299-1
- Silvia, P. J., Martin, C., & Nusbaum, E. C. (2009). A snapshot of creativity: Evaluating a quick and simple method for assessing divergent thinking. *Thinking Skills and Creativity*, 4, 79-85. doi: 10.1016/j.tsc.2009.06.005
- Silvia, P. J., Winterstein, B. P., & Willse, J. T. (2008). The madness to our method: Some thoughts on divergent thinking. *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 109-114. doi: 10.1037/1931-3896.2.2.109
- Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity and the Arts*, 2(2), 68-85. doi: 10.1037/1931-3896.2.2.68
- Torrance, E. P. (1966). *Torrance tests of creative thinking*. Lexington: Personnel Press.
- Treffinger, D. J., Schoonover, P. F., & Selby, E. C. (2013). *Educating for creativity & innovation*. Texas: Prufrock Press.

- Vilarinho-Rezende, D., Borges, C. N., Fleith, D. S., & Joly, M. C. R. A. (2016). Relação entre tecnologias da informação e comunicação e criatividade: Revisão da literatura. *Psicologia: Ciência e Profissão*, *36*(4), 877-892. doi: 10.1590/1982-3703001342014
- Wallach, M. A., & Kogan, N. A. (1965). A new look at the creativity-intelligence distinction. *Journal of Personality*, *33*, 348-369. doi: 10.1111/j.1467-6494.1965.tb01391.x

recebido em agosto de 2018
aprovado em dezembro de 2019

Sobre a autora

Tatiana de Cássia Nakano é Doutora em Psicologia pela Pontifícia Universidade Católica de Campinas. Docente do curso de Pós-graduação *Stricto Sensu* em Psicologia da PUC-Campinas.