

Approaches to Measuring Creativity: A Systematic Literature Review*

Sameh Said-Metwaly

University of Leuven, Belgium

E-mail address: sameh.metwaly@kuleuven.be

Eva Kyndt

University of Leuven, Belgium

E-mail address: eva.kyndt@kuleuven.be

Wim Van den Noortgate

University of Leuven, Belgium

E-mail address: wim.vandennoortgate@kuleuven.be

ARTICLE INFO

Keywords:

Creativity measurement
Creativity assessment
Creativity instruments
Creativity ratings
Creativity testing

Article history:

Received 08 September 2017

Received in revised form 05 December 2017

Accepted 06 December 2017

ISSN: 2354-0036

DOI: 10.1515/ctra-2017-0013

ABSTRACT

This paper presents a review of the literature on the measurement of creativity. Creativity definitions are discussed as a starting point for understanding the nature of this construct. The four major approaches to measuring creativity (process, person, product and press) are reviewed, pointing out commonly used instruments as well as the advantages and weaknesses of each approach. This review reveals that the measurement of creativity is an unsettled issue, and that the existing instruments purporting to measure creativity suffer from serious conceptual and psychometric shortcomings. Research gaps and suggestions for future research are discussed.

INTRODUCTION

Creativity is considered as a precious asset for solving individual, organizational and social problems and achieving sustainable development (Barbot, Besançon, & Lubart, 2015; Lubart, Zenasni, & Barbot, 2013; Zeng, Proctor, & Salvendy, 2011). Accordingly, understanding the nature and determinants of creativity is of great importance to enhance its development. However, although much research has been devoted to such understanding, some controversial issues have not yet been settled.

One of the most controversial issues in creativity research is the measurement of this construct (Baer & McKool, 2009; Barbot, Besançon, & Lubart, 2011; Blamiresa & Peterson, 2014; Kaufman, Baer, Cole, & Sexton, 2008; Mouchiroud & Lubart, 2001; Park, Chun, & Lee, 2016; Zeng et al., 2011). The growing body of creativity research has devoted considerable attention to the development of numerous instruments assessing dif-

* This research was supported by a grant from the Government of Egypt

ferent aspects of creativity, ranging from cognitive processes to environmental factors (Kaufman, Lee, Baer, & Lee, 2007; Krumm, Lemos, & Filippetti, 2014; Park et al., 2016). The most prominent characteristic of this body of instruments is their diversity, indicating the complexity of defining and measuring creativity (Hocevar, 1981). In addition, these instruments have limitations, such as measuring trivial aspects of creativity or lacking adequate psychometric properties (Baer & McKool, 2009; Kaufman et al., 2007). Hence, researchers seeking to measure creativity face a number of unsolved issues, including how creativity should be defined, what particular aspects of this construct need to be valued and measured, and what the most appropriate instruments are, for measuring these aspects (Blamiresa & Peterson, 2014; Treffinger, Young, Selby, & Shepardson, 2002). In this regard, this paper aims to shed light on these issues through answering these questions:

1. How is creativity defined and described in the literature?
2. What are the approaches to measuring creativity?
3. What are the most commonly used instruments in each approach?
4. What are the advantages and weaknesses of the existing approaches for measuring creativity?

We try to answer these questions through systematically reviewing the existing literature. This literature review seeks to provide researchers with an up-to-date overview of the current state of creativity measurement, including the existing measurement approaches and the psychometric properties of commonly used instruments, which could help to increase their awareness of the existing research gaps and questions that need to be approached in future research. Broadly speaking, our review might help to provide valuable insights into the field of creativity and its measurement and contribute to setting the research agenda for further validation of the existing creativity instruments.

METHOD

The present review includes creativity literature published in books and journals up to December 31st, 2016. The search process consisted of the following four steps: First, the following databases were searched: ERIC, Google Scholar, JSTOR, PsycINFO and Web of Science using the following search string: (“creativity” OR “creative thinking” OR “creative performance” OR “creative ability” OR “creative potential”) AND (“measurement” OR “assessment” OR “evaluation” OR “testing”). Second, the reference lists of the papers identified in the first step were reviewed for other relevant references (i.e. “backward search”). Third, more recent references were retrieved by searching databases for papers that referred to the previously identified papers in steps 1 and 2 in their citations (i.e.

“forward search”). Fourth, the following key journals in creativity were hand-searched: *Creativity Research Journal*, *Gifted Child Quarterly*, *Psychology of Aesthetics*, *Creativity, and the Arts*, *The Journal of Creative Behavior* and *Thinking Skills and Creativity*.

The papers identified using the search process were first screened for their relevance on the basis of their titles and abstracts. The remaining papers were included if they met the following two criteria: (1) addressed or discussed the measurement of creativity, and (2) reported a development or validation of instruments measuring creativity. Moreover, we only included (1) journal articles, conference papers, or dissertations (2) that were written in English, and for which (3) the full text was available. Studies using qualitative approaches for measuring creativity and intervention studies were excluded. We further included creativity instruments that met the following criteria: (1) were published in English (2) were developed to measure one of the aspects of creativity identified in the review framework, and (3) were employed in at least five published empirical studies.

The analysis of the data was informed by the review’s objective and the questions presented above. A thematic approach was employed to analyze the final list of included papers. The analysis involved categorizing the content of the included papers around the following themes: definitions of creativity, approaches to measuring creativity and the advantages and weaknesses of the measurement approaches. Data extracted from empirical papers included research objectives, the measurement approach and instrument used to measure creativity, data analysis method and psychometric properties findings for the creativity instruments (i.e. reliability and validity).

RESULTS

From the 2,064 papers identified by the search process, 221 papers were selected based on screening titles and abstracts. Among these, 152 papers met the inclusion criteria. The 152 included papers addressed the measurement of creativity and significant issues related to this measurement. Four distinct approaches to measuring creativity (process, person, product and press), in addition to the most commonly used instruments in each approach were identified.

In the following, we first discuss creativity definitions, pointing to the different categories of these definitions. Then, we describe the approaches to measuring creativity and the advantages and weaknesses of each of these approaches, with an emphasis on the psychometric properties of the most common instruments used in each approach.

Defining creativity

Creativity has proven, over the years, to be difficult to define and measure due to its complex and multidimensional nature (Barbot et al., 2011; Batey & Furnham, 2006; Cropley, 2000; Runco, 2004, 2007; Treffinger et al., 2002). Treffinger (1996) reviewed the creativi-

ty literature and presented more than 100 different definitions for this concept. Despite these different definitions, the majority of creativity studies tend to employ only a few of these definitions, whereas other studies avoid providing a definition of this construct at all (Kaufman, Plucker, & Russell, 2012; Plucker & Makel, 2010). Furthermore, researchers and educators may use the term creativity to refer to entirely different aspects, including cognitive processes, personal characteristics and past experiences (Treffinger et al., 2002). In addition, researchers sometimes use terms such as innovation, invention, imagination, talent, giftedness and intelligence interchangeably with creativity.

In general, definitions of creativity typically reflect at least one of four different perspectives: cognitive processes associated with creativity (later in this paper referred to as 'process'), personal characteristics of creative individuals ('person'), creative products or outcomes ('product') and the interaction between the creative individual and the context or environment ('press') (Couger, Higgins, & McIntyre, 1993; Horn & Salvendy, 2006; Rhodes, 1961; Thompson & Lordan, 1999; Zeng et al., 2011).

With regard to the process perspective, Torrance (1977), as a pioneer in creativity research, defined creativity as the process of perceiving problems or gaps in knowledge, developing hypotheses or propositions, testing and validating hypotheses and finally sharing the results. Similarly, Mednick (1962) proposed that creativity involves the process of bringing associative elements together into new combinations to meet the task requirements. Guilford (1950) suggested some factors for interpreting variations in creativity including sensitivity to problems, fluency, flexibility, originality, synthesizing, analyzing, reorganizing or redefining, complexity and evaluating. In his Structure-of-Intellect (SOI) Model, Guilford (1975) considered creativity as a form of problem solving and distinguished between two types of cognitive operations: divergent production and convergent production. Divergent production is a broad search used in open problems to generate logical answers or alternatives, whereas convergent production is a focused search that leads to the generation of a specific logical imperative for a problem, in which a particular answer is required. Guilford (1975) considered divergent production process to be more relevant to successful creative thinking.

Focusing on the person perspective, a wide array of personal characteristics and traits have been suggested as being associated with creativity including attraction to complexity, high energy, behavioural flexibility, intuition, emotional variability, self-esteem, risk taking, perseverance, independence, introversion, social poise and tolerance to ambiguity (Barron & Harrington, 1981; Feist, 1998; James & Asmus, 2000-2001; Runco, 2007). However, having such traits does not actually guarantee the occurrence of creative

achievement, the effect of intrinsic motivation still remains (Amabile, 1983). In other words, personality may be seen as related to the motivation to be creative rather than to creativity itself, with both of these being necessary for creative achievement (James & Asmus, 2000-2001). Task motivation is one of three key components in Amabile's (1983, 1988, 1996) componential model of creativity that are necessary for creative performance, together with domain-relevant skills (including knowledge about the domain, technical skills and domain-related talent) and creativity-relevant skills (including personality characteristics and cognitive styles).

By turning the focus of defining creativity towards the creative products, Khatena and Torrance (1973) defined creativity as constructing or organizing ideas, thoughts and feelings into unusual and associative bonds using imagination power. Gardner (1993) stated that creative individuals are able to solve problems, model products, or define new questions in a novel but acceptable way in a particular cultural context. Creativity is also seen as the ability to produce or design something that is original, adaptive with regard to task constraints, of high quality (Kaufman & Sternberg, 2007; Lubart & Guignard, 2004; Sternberg & Lubart, 1999), useful, beautiful and novel (Feist, 1998; Mumford, 2003; Ursyn, 2014).

Finally, regarding the press perspective, that is, the interaction between the creative person and the environment or climate, McLaren (1993) stated that creativity could not be fully understood through human endeavour without taking into account its socio-moral context and intent (James, Clark, & Cropanzano, 1999). Investigating the environment for creativity therefore requires that all the factors that promote or inhibit creativity should be taken into consideration (Thompson & Lordan, 1999). In the componential model of organizational innovation and creativity, Amabile (1988) proposed three broad environmental factors related to creativity: organizational motivation or orientation to innovate, available resources and management practices. Geis (1988) identified five factors to ensure a creative environment: a secure environment with minimum administrative or financial intervention, an organizational culture that makes it easy for people to create and discover independently, rewards for performance to support intrinsic motivation, managerial willingness to take risks in the targeted areas of creativity and providing training to enhance creativity. Several studies have indicated the impact of climate or environment variables on creative achievement (e.g. Couger et al., 1993; Paramithaa & Indarti, 2014), particularly with respect to the initial exploratory stages of creative endeavours in which individuals' need for approval and support plays an important role in motivating their further efforts (Abbey & Dickson, 1983).

Despite these different perspectives in defining creativity, some aspects are shared

by many researchers. Researchers generally agree that creativity involves the production of novel and useful responses (Batey, 2012; Mayer, 1999; Mumford, 2003; Runco & Jaeger, 2012). These two characteristics, novelty and usefulness, are widely mentioned in most definitions of creativity (Zeng, Proctor, & Salvendy, 2009), although there is still some debate about the definitions of these two terms (Batey, 2012; Batey & Furnham, 2006; Runco & Jaeger, 2012). Another area of consensus is that creativity is regarded as a multifaceted phenomenon that involves cognitive, personality and environmental components (Batey & Furnham, 2006; Lemons, 2011; Runco, 2004). As Harrington (1990, p.150) asserted “Creativity does not “reside” in any single cognitive or personality process, does not occur at any single point in time, does not “happen” at any particular place, and is not the product of any single individual”.

Approaches to measuring creativity

The growth of creativity research has been associated with looking for dependable instruments to measure it. However, over the last number of decades measuring creativity has proven to be a challenging task (Baer & McKool, 2009; Batey, Chamorro-Premuzic, & Furnham, 2010; Kaufman et al., 2008). Much of the ambiguity surrounding the measurement of creativity is attributed to the lack of consensus among researchers on its definition (Batey, 2012; Piffer, 2012). Based on different definitions of creativity, researchers have developed different instruments for measuring it (Batey, 2012; Belcher, Rubovits, & Di Meo, 1981; Horn & Salvendy, 2006). Each instrument reflects the conception of its developer regarding the nature of creativity (Treffinger, Renzulli, & Feldhusen, 1971). These instruments are typically classified into four approaches representing the four main categories of creativity definitions: process, product, person and press (Barbot et al., 2011; Couger et al., 1993; Fishkin & Johnson, 1998; Horn & Salvendy, 2006; Rhodes, 1961; Thompson & Lordan, 1999). Table 1 gives an overview of the approaches to measuring creativity. Based on our analysis, the process approach was the most common approach used to measure creativity (52.58% of our sample of 152 studies). The person approach (28.87%) and the product approach (14.43%) were the second and third most commonly used approaches. The press approach was the least common among these four approaches (4.12%). In total, 18 instruments were found (process $n = 4$, person $n = 6$, product $n = 1$ and press $n = 7$; see Table 1).

Table 1
Approaches of Measuring Creativity

Approach	Focus	Conception of Creativity	Instruments	Advantages	Weaknesses
Process	Creative processes or skills associated with creativity	Domain-general	WKCT (Wallach & Kogan, 1965); TTCT (Torrance, 1966, 2008); SOI (Guilford, 1967); CAP (Williams, 1980)	Widespread utility High reliability Standardized criteria for interpreting scores	Limited scope of measurement conflicting evidence for validity Bias due to scoring and sample size
Person	Personality traits or creative achievements	Domain-general or domain-specific	HDYT (Davis & Subkoviak, 1975); CPS (Gough, 1979); HCAY (Raudsepp, 1981); CBI (Hocevar, 1979c); CAQ (Carson et al., 2005); BICB (Batey, 2007)	Ease of use High reliability Standardized criteria for interpreting scores	Limited scope of measurement Low validity of self-reports Bias due to self-reporting Neglect of differences in creative personality across domains Low sensitivity to training Skewed scores
Product	Creative products	Domain-specific	CAT (Amabile, 1982)	Similar to evaluating creativity in real-life High reliability High validity	Limited scope of measurement Difficulty in selecting judges Bias due to judges Expensive and time-consuming Lack of standardized criteria
Press	Work environment or climate	Domain-general	SSSI (Siegel & Kaemmerer, 1978); CUCEI (Fraser et al., 1986); WES (Moos, 1986); WEI (Amabile & Gryskiewicz, 1989); KEYS (Amabile et al., 1996); TCI (Anderson & West, 1998); SOQ (Isaksen et al., 1999)	Explore whether a work environment is supportive or inhibitive of creativity Evaluate the environmental improvement attempts and corrective actions	Limited scope of measurement Lack of research-based evidence Debate about “climate” meaning and measurement level Individual differences in conception of climate

Process approach. The process approach in creativity measurement focuses on the specific cognitive processes and structures that are conducive to creative production (Barbot et al., 2011; Finke, Ward, & Smith, 1992). Divergent thinking tests have been most widely used for measuring creative processes or creativity-relevant skills (Kaufman et al., 2007; Plucker, Qian, & Wang, 2011; Zeng et al., 2011). Examples of these tests include the Wallach-Kogan Creativity Tests (WKCT; Wallach & Kogan, 1965), the Torrance Tests of Creative Thinking (TTCT; Torrance, 1966, 2008), the Structure of the Intellect Divergent Production Tests (SOI, Guilford, 1967) and the Creativity Assessment Packet (CAP; Williams, 1980). These tests, which are also known as measures of ideational fluency, include open or ill-structured problems that require individuals to generate as many responses as possible, which are then scored to capture fluency (number of responses), originality (statistical rarity), flexibility (number of different categories) and elaboration (amount of detail) (Jauk, Benedek, & Neubauer, 2014; Plucker et al., 2011). Hence, the focal point of divergent thinking tests is not only to consider the amount of responses, but also the quality of these responses (Silvia et al., 2008).

Despite their prevalence, divergent thinking tests and their effectiveness in measuring creativity have long been a subject of debate. Much research has been conducted to examine the psychometric properties of these tests. While evidence of the *reliability* of divergent thinking tests has been obtained (e.g. Cropley, 2000; Cropley & Maslany, 1969b; Hocevar, 1979b; Lemons, 2011; Torrance, 2008; Torrance & Haensly, 2003), many questions have been raised about their *validity* in measuring creativity (Baer, 2016; Cropley, 2000; Lemons, 2011; Hennessey & Amabile, 2010). Regarding their *construct validity*, little evidence is available. First, studies have yielded contradictory results regarding the latent structure of divergent thinking tests. For example, some studies have shown that the scores for the TTCT primarily reflect one general creativity factor (e.g. Clapham, 1998; Dixon, 1979; Heausler & Thompson, 1988; Hocevar, 1979d); other studies have found more than one factor (e.g. Almeida, Prieto, Ferrando, Oliveira, & Ferrándiz, 2008; Clapham, 2004; Kim, 2006b; Kim, Cramond, & Bandalos, 2006; Krumm, Aranguren, Filippetti, & Lemos, 2014; Krumm, Filippetti, Lemos, Koval, & Balabanian, 2016; Krumm, Lemos, & Filippetti, 2014). Similarly, contrasting results have been reported about the factor structure of the SOI. Some studies have indicated the presence of a general factor accounted for much of the total variance of the SOI tests due to the high intercorrelations (exceeding .90 in most cases) among different factors (e.g. Hocevar, 1979a, 1979b; Khattab, Michael, & Hocevar, 1982; Mace, Michael, & Hocevar, 1985), while other studies have revealed the multidimensionality of the SOI (e.g. Bachelor & Michael, 1991; Bache-

lor, Michael, & Kim, 1994; Michael & Bachelor, 1990; Ulosevich, Michael, & Bachelor, 1991). More interestingly, Bachelor (1989) concluded that neither the general single factor models nor the high-dimensionality models could provide a reasonable representation of the SOI data.

Second, several studies have failed to demonstrate strong *discriminant validity* for the TTCT (e.g. Clapham, 1998; Dixon, 1979; Heausler & Thompson, 1988; Hocevar & Michael, 1979; Kim et al., 2006; Mouchiroud & Lubart, 2001; Runco & Mraz, 1992; Silvia et al., 2008) as well as the SOI dimensions (e.g. Hocevar, 1979a, 1979b; Hocevar & Michael, 1979). The results of these studies indicated that originality and fluency scores derived from divergent thinking tests are highly correlated, with observed correlation coefficients that often ranged between .79 and .86 (e.g. Kim et al., 2006), which raises doubt about whether the fluency dimension contributes significantly to explaining any variance beyond the fluency dimension itself (Silvia et al., 2008). Some suggestions regarding originality scoring were made to decrease the confusion among originality and fluency; however, it was found that when fluency subscores were partialled out, the originality subscores often became unreliable (e.g. Hocevar, 1979a, 1979b, 1979d; Hocevar & Michael, 1979; Runco & Mraz, 1992; Runco, Okuda, & Thurston, 1987; Silvia et al., 2008). Another problematic psychometric property related to discriminant validity is that originality scores could depend on the sample size (Silvia, Martin, & Nusbaum, 2009). The probability that an original response will appear in individuals' responses is a function of the sample size; as the sample size increases, the likelihood that a response is repeated increases, so unique responses become less prevalent (Silvia et al., 2008). It thus becomes less possible to distinguish original responses in large samples, which may lead to inconsistent results among creativity studies due to the influence of different sample sizes.

With regard to *content validity*, divergent thinking tests have been frequently criticized because of their limited ability to represent all facets of creativity as a multidimensional construct. There are many possible reasons to account for the low content validity of divergent thinking tests. First, the key assumption behind the development of divergent thinking tests is that creativity represents a general and identical construct across all domains, and hence the items included in these tests are relatively content-free and domain-independent (Diakidoy & Spanoudis, 2002; Weisberg, 2006). However, a growing body of research indicates that creativity may be best conceived as a domain-specific construct and that general skills or processes only have a limited contribution to creative achievement (e.g. Baer, 1991, 1994a, 1994b; Han, 2003; Han & Marvin, 2002; Palmiero, Nori, Aloisi, Ferrara, & Piccardi, 2015; Reiter-Palmon, Robinson-Morrall, Kaufman, & Santo, 2012). Second, it has been suggested that there are four basic phases in the creative

process: problem analysis, ideation, evaluation and implementation (Lubart, 2001; Zeng et al., 2009, 2011). Nevertheless, divergent thinking tests emphasize only one of these phases, ideation, neglecting the connectedness among the four phases in producing creative solutions (Zeng et al., 2011). Third, performance on divergent thinking tests is restricted to the behaviours being aroused through the given items (Fishkin & Johnson, 1998). As a result, this performance might not perfectly exemplify the actual creative production fulfilled by creative people. In addition, this performance might lack the high personal commitment that characterizes self-initiated products (Fishkin & Johnson, 1998). In other words, the responses to these items do not assure the success of individuals in employing these skills in real-life situations, or their success in determining and implementing the promising ideas from different ideas generated. Consequently, many researchers have pointed out that performance on divergent thinking tests should not be treated as fully representative of an individual's entire creativity (Baer, 2011, 2016; Batey & Furnham, 2006; Fishkin & Johnson, 1998; Kaufman et al., 2007; Kaufman et al., 2012; Kim, 2006a; Lubart & Guignard, 2004; Runco, 1993, 2008; Runco & Acar, 2012; Torrance & Haensly, 2003; Treffinger et al., 1971; Zeng et al., 2011).

With regard to the *criterion-related validity* of divergent thinking tests, both concurrent validity and predictive validity have been investigated. In relation to the *concurrent validity* of divergent thinking tests, research has been carried out to examine the extent to which performance on these tests is related to the actual creative achievement. This relationship has been examined in a wide range of creative achievement domains, including mathematics, science, language, real-life problem solving, art, literature, leadership and craft. Some studies in this area found a positive association between divergent thinking and creative achievement scores (e.g. Clapham, 2004; González, Campos, & Pérez, 1997; Jauk et al., 2014; King, Walker, & Broyles, 1996; Runco, 1984; Runco & Okuda, 1988; Vincent, Decker, & Mumford, 2002; Wechsler, 2006), while others found a negative or no association (e.g. Brougher & Rantanen, 2009; Han, 2003; Han & Marvin, 2002). Other studies indicated that the positive association is restricted to certain subpopulations (i.e. gifted vs. non-gifted), certain creative achievement criteria (i.e. quantitative vs. qualitative) or certain domains (e.g. writing, music, art or crafts), but with a lack of consensus on any of the alternatives (e.g. Davidovitch & Milgram, 2006; Okuda, Runco, & Berger, 1991; Runco, 1986, 1987).

These mixed results for concurrent validity could be attributed to the different instruments used in measuring creative thinking. Strong evidence of concurrent validity was found in studies that used the TTCT (e.g. Clapham, 2004; González et al., 1997; Jauk et al., 2014; King et al., 1996; Wechsler, 2006), whereas weak evidence was found in stud-

ies that used other tests such as the WKCT (e.g. Han, 2003; Han & Marvin, 2002; Okuda et al., 1991; Runco, 1986, 1987). The items in the TTCT might be more likely than the items in the WKCT to elicit processes and responses that are characteristic of a person's ability to apply these skills in real-world situations. Unfortunately, there are no comparable data for the concurrent validity of different divergent thinking tests to validate this hypothesis. In addition, the limited concurrent validity of divergent thinking tests may be due to the fact that there are many factors, besides the divergent thinking ability, that contribute to creative achievement, including cognitive, motivational, social and environmental factors (Ericsson, 1999; Wechsler, 2006). Divergent thinking skills may be an essential element of creativity, but alone cannot explain the variation in creative achievement (Baer, 2016; Batey & Furnham, 2006; Runco, 1993). Moreover, these studies did not take the influence of expertise on the relationship between divergent thinking and creative achievement into account. Research has indicated that most of the individual differences in creative achievement are related to the accumulated amount of expertise (Ericsson, Krampe, & Tesch-Römer, 1993; Weisberg, 2006). It has been suggested that expert performers in different domains have developed mental representations and skills that help them to select and monitor the best courses of action under situations requiring creativity (Ericsson, 1999). Thus, the concurrent validity of divergent thinking tests deserves further examination in light of direct and indirect effects between divergent thinking and expertise.

With respect to *predictive validity*, there is currently a great deal of controversy about whether divergent thinking tests could be considered as valid predictors of future creative achievement. The results of Torrance's studies (Torrance, 1969, 1972a, 1972b, 1980, 1981a, 1981b), that followed individuals from 7 to 22 years of age, showed that the TTCT scores were significantly better predictors of creative achievement than intelligence scores. More recent studies analyzed Torrance's longitudinal data collected over a longer period of time (from 22 to 50 years) using regression and structural equation modelling (SEM) and asserted the greater predictive power of TTCT over intelligence scores (e.g. Cramond, Matthews-Morgan, Bandalos, & Zuo, 2005; Plucker, 1999a; Runco, Millar, Acar, & Cramond, 2010). Milgram and Hong (1993) examined the predictive validity of the WKCT by utilizing data collected over 18 years and found that these tests were better predictors of adult life accomplishment than intelligence or school grades. Cropley (1972) administered the SOI to grade seven students and assessed nonacademic achievement in four areas (art, drama, literature and music) five years later. The results indicated that there was a substantial longitudinal relationship between total SOI scores and nonacademic achievement, and adding intelligence scores to the predictors improved their predictive power only slightly. Furthermore, meta-analyses (Gajda, Karwowski, & Beghetto,

2016; Kim, 2008) reported that divergent thinking scores accounted for more variance in creative achievement than did intelligence scores. However, other researchers failed to provide evidence supporting the predictive power of divergent thinking tests (e.g. Getzels & Csikszentmihalyi 1976; Hocevar, 1980; Kogan & Pankove, 1974; Skager, Klein, & Schultz, 1967).

Several explanations have been proposed for the lack of consistency among studies that addressed the predictive validity of divergent thinking tests. One explanation is the short period over which longitudinal studies of divergent thinking are conducted (Hocevar & Bachelor, 1989; Plucker, 1999a; Torrance, 1972b). Torrance (1972a) suggested that an appropriate period for follow-up studies of adult creative achievement is at least twelve years following high school graduation. This suggestion is in agreement with the results of longitudinal studies that provided evidence supporting the predictive validity of divergent thinking tests (e.g. Cramond et al., 2005; Plucker, 1999a; Runco et al., 2010; Milgram & Hong 1993; Torrance, 1972a, 1972b, 1980, 1981a, 1981b). A second possible explanation may be the fact that irrelevant outcome criteria have been used (Baer, 2011; 2016; Han, 2003; Plucker & Runco, 1998; Runco, 1986; Weisberg, 2006). Criteria frequently used in creativity longitudinal studies are self-report measures that ask subjects whether they have engaged in particular activities (e.g. writing articles, poems, stories or songs, participating in political campaigns, inventing devices, or winning prizes or grants) (see, for example, Hocevar, 1980; Holland & Nichols, 1964; Torrance, 1969). Subjects' responses on such self-report measures might over-report creativity (Baer, 2011, 2016) or subjects may lack insight into their own creative accomplishments (Kaufman, Evans, & Baer, 2010). Moreover, these measures only consider the quantitative component of creative accomplishment. This creates quite a problem, as the qualitative component of creative achievement is possibly more important in most of real-life situations (Harrington, 1975; Runco, 1986, 1987), and further, these two components are highly correlated with each other (Cramond et al., 2005). This could explain the low predictive validity obtained in longitudinal studies that only included quantitative indices of creative achievement (e.g. Han, 2003; Hocevar, 1980; Kogan & Pankove, 1974). Besides, such self-reported activities do not necessarily require high levels of divergent thinking ability, and more intelligent people are expected to be able to engage in such activities (Weisberg, 2006). Creative people might do these activities more creatively, but not necessarily more repeatedly (Baer, 2016). To complicate matters further, recent studies (Primi, 2014; Silvia, 2015) have stated that divergent thinking and intelligence are more closely related than initially thought. If this is generally the case, then the variance of creative achievement may largely be attributed to differences in intelligence, not to differences in divergent thinking ability

(Weisberg, 2006). Baer (1994d) stressed that divergent thinking tests may not provide additional information on creative achievement when compared to intelligence scores due to the overlap between what divergent thinking tests and intelligence tests measure.

Furthermore, the observed limited predictive validity may be attributed to the use of inadequate statistical methods to analyze longitudinal data (Hocevar & Bachelor, 1989; Plucker, 1999a; Plucker & Renzulli, 1999). For instance, most statistical analyses employed in Torrance's (1972a) study were simple correlation coefficients that do not control for the effect of other external variables. Plucker's (1999a) reanalysis of these data using SEM yielded more positive results regarding the predictive validity of the TTCT. More importantly, such a statistical technique improves conventional analysis by providing researchers with opportunities to investigate relationships between latent variables, while considering potential measurement errors among targeted variables. Last but not least, the limited predictive validity of divergent thinking tests may be a result of reliance on the domain-generality rather than the domain-specificity of creativity. In other words, previous studies might have assumed that divergent thinking ability includes various skills that fit all domains, and an individual who scores highly on divergent thinking tests would be expected to exhibit creativity in any domain. Thus, these studies did not take the creativity-relevant skills in each domain into account. Runco (1987) pointed to the need to consider the domain-specificity when examining the predictive power of divergent thinking tests. Following this interpretation, it could be argued that each of the divergent thinking skills plays a role in one specific performance domain or only in specific tasks within this domain (Baer, 1993; Runco et al., 2010). Torrance (1972a) found that creative achievements were more easily predicted by divergent thinking tests in specific domains (writing, science, medicine and leadership) than in other domains (music, the visual arts, business and industry). Plucker's (1999a) study indicated that a verbal (but not figural) divergent thinking test was a better predictor of adult craft creative achievement than intelligence. Plucker explained this result by arguing that most creative achievements may require high levels of linguistic talent, in contrast to spatial or problem solving talents. In addition, Cramond et al. (2005) reported that originality and fluency scores were the best predictors of quantitative creative achievement, while flexibility and originality were the best predictors of qualitative creative achievement. In conclusion, Barron and Harrington (1981) went on to suggest that, under certain conditions, some divergent thinking tests scored according to definite criteria could predict creative achievement in specific domains. With this in mind, researchers need to address all of these variables when making decisions about the ability of divergent thinking to predict creative achievement.

Person approach. This approach for measuring creativity relies heavily on using self-report questionnaires that enquire about personality traits related to creativity. Instruments in this area were designed by investigating the characteristics or interests of individuals who produced creative achievements. This investigation, as discussed earlier, generated a set of distinctive personality traits for creative individuals, including attraction to complexity, high energy, behavioural flexibility, intuition, emotional variability, self-esteem, risk taking, perseverance, independence, introversion, social poise and tolerance to ambiguity (Barron & Harrington, 1981; Feist, 1998; James & Asmus, 2000-2001; Runco, 2007). An individual who exhibits these characteristics is more likely to act creatively than an individual without these characteristics. Several questionnaires have been developed in this area, such as How Do You Think? (HDYT; Davis & Subkoviak, 1975), the Creative Personality Scale (CPS; Gough, 1979) and How Creative Are You? (HCAY; Raudsepp, 1981). Other self-report questionnaires ask individuals about their past creative production instead of personality traits, based on the assumption that past accomplishments can provide a dependable indication of subsequent creative accomplishments (Clapham, 2004; Colangelo, Kerr, Hallowell, Huesman, & Gaeth, 1992). Such questionnaires include the Creative Behavior Inventory (CBI; Hocevar, 1979c), the Creative Achievement Questionnaire (CAQ; Carson, Peterson, & Higgins, 2005) and the Biographical Inventory of Creative Behaviors (BICB; Batey, 2007).

Although this approach for measuring creativity has advantages of ease of use, standardized administration and scoring procedures (Runco, 2007) and high levels of reliability (see Carson et al., 2005; Davis, 1975; Gough, 1979; Hocevar, 1980; Silvia, Wigert, Reiter-Palmon, & Kaufman, 2012), it suffers from significant limitations. First, this approach mainly pays attention to personality traits, as it assumes that these could be regarded as sufficient evidence of creative ability (Barbot et al., 2011; Miller, 2014; Thompson & Lordan, 1999). It seems that this approach closes its eyes to the notion that creativity is a multidimensional and complex construct. In addition, no one is expected to have all of the traits reported in the creativity literature. Even if persons have such traits, they may not produce creative accomplishments due to their lack of abilities, attitudes, motivation or supporting environment.

Second, a great deal of controversy exists about the validity of self-report measures of creativity. Previous studies have documented a significant relationship between scores obtained from different self-report measures of creativity (e.g. Fleenor & Taylor, 1994; Furnham, Batey, Anand, & Manfield, 2008; Goldsmith & Matherly, 1988; Reiter-Palmon et al., 2012). However, this is in contrast with the results of studies that examined the rela-

relationship between these measures and actual creative performance. Priest (2006) revealed that the self-ratings of undergraduate students' of their musical compositions were significantly different from experts' ratings of these compositions. Similarly, Kaufman et al. (2010) asked fourth-grade students to rate their own creativity in four domains (science, maths, writing and art). The correlations between self-rated creativity and experts' ratings of students' work in these domains were non-significant and ranged from -.22 to .07. Reiter-Palmon et al. (2012) examined the relationship between self-perceptions of creativity, creative personality, creative self-efficacy and creative problem solving (participants' solutions to problems were evaluated by three trained raters). Although significant correlations were found between self-perceptions and self-report measures of creativity, correlations with measures of creative problem solving were non-significant and sometimes even negative. These results raise doubts about the validity of the information obtained from self-report measures of creativity.

Third, data based on self-reports of creativity, like other self-reports, are subject to potential bias that endangers their validity (Kaufman et al., 2007). Individuals may distort their responses on self-reports, intentionally or unintentionally, for many reasons. These reasons include social desirability, consistency motive, mood state and implicit theories. Social desirability refers to the tendency of individuals to report about themselves in a favourable manner, regardless of their actual behaviour (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). As in the case of measuring creativity, the threat of social desirability increases in situations where individuals feel that the results of evaluation matter to them personally (Kaufman et al., 2007; Ng & Feldman, 2012). The consistency motive refers to the tendency of individuals to appear consistent in their responses to self-report items (Podsakoff et al., 2003). This is most likely to exist when individuals are asked to respond to items about their past attitudes or behaviours (Podsakoff et al., 2003). Such a tendency may have a negative influence on the results of self-reports of creativity because respondents may try to keep consistency in their responses to items, which may result in artificially high relationships that may not exist in real-life settings (Ng & Feldman, 2012). Mood state can be defined as a temporary state or feeling that causes individuals to see themselves and the surrounding world in positive or negative terms (Podsakoff et al., 2003). Whether individuals' mood is positive or negative, it may influence their responses on self-reports of creativity in the same manner (Ng & Feldman, 2012). Finally, implicit theories refer to personal conceptions (informal theories) underlying in the minds of individuals regarding a particular phenomenon (Sternberg, Conway, Ketron, & Bernstein, 1981). Because of diverse implicit theories, individuals may explain the items differently

and subjectively define the anchors of their self-reports (Biernat, 2003). Likewise, individuals seem to have different conceptions and notions when they use the term creativity and rely on these conceptions in judging their own creativity and the creativity of others (Kaufman & Baer, 2004; Sternberg, 1985). Runco and Johnson (1993) compared parents' and teachers' implicit theories of creativity. The results revealed that although some similar traits were selected to describe creativity in these two groups, other traits were only selected by either parents (e.g. impulsive, resourceful and enterprising) or teachers (e.g. cheerful, friendly and friendly). Similarly, Westby and Dawson (1995) showed that the traits identified by teachers as characteristics of creative children were significantly different from those presented in the creativity literature. More importantly, teachers' conceptions of the characteristics of creative children were more related to the characteristics that make them easy to be handled in the classroom (e.g. responsible, good-natured, reliable and sincere). Furthermore, other studies found differences in implicit theories of creativity between experts and non-experts (Runco & Bahleda, 1986), experts in different domains (Sternberg, 1985) or experts in the same domain (Gluck, Ernst, & Unger, 2002). Taken together, these results suggest that when individuals evaluate their own creativity, they may do so with reference to their own implicit theories of creativity that are more likely to differ, either from each other, or even from explicit theories of creativity. Consequently, individuals' evaluations of their own creativity cannot be directly compared, as the conceptions on which these evaluations are built are related to distinctive referents (Biernat, 2003).

Fourth, most self-reports of creativity (e.g. HDYT, CPS and HCAY) adopt a general perspective in measuring creativity, without any indication of those particular personality traits attached to each domain (Kaufman, 2012; Kaufman & Baer, 2004). Runco (2007) emphasized that a unique creative personality does not exist, and creative personality is different from domain to domain, and possibly from person to person. This is in line with the results of Feist's (1998, 1999) meta-analysis of differences in personality traits among creative scientists and creative artists. Feist (1998, 1999) concluded that, although these two groups generally shared some personality traits (e.g. openness to experience, conscientiousness, hostility and impulsivity), a number of traits were more characteristic within each group. Creative scientists tended to be self-confident, autonomous and arrogant, while creative artists tended to be anxious, imaginative and emotionally sensitive. Reanalyzing the self-rated creative achievement data in different domains obtained from three previous studies, Plucker (1999b) found a general factor that accounted for more than 40% of the variance. Plucker (1999b) pointed out that the variance unexplained by the general factor may be attributed to the domain-specific factors. He also suggested that if the participants were older than undergraduate students, the contribution of these specific

factors might increase. Likewise, Kaufman et al. (2010) found that the correlations between students' self-reported creativity in four domains (science, maths, writing and art) were low (less than .30) and non-significant in most cases. Based on these results, researchers should be cautious in using self-report measures of creativity, especially when they do not reflect domain-relevant aspects of creativity.

Fifth, self-reports of creativity are less sensitive to the effects of creativity training than performance measures of creativity (Fishkin & Johnson, 1998). In their meta-analysis on the effectiveness of gifted education programs, Vaughn, Feldhusen, and Asher (1991) concluded that the effects of these programs on self-report measures of creativity were limited when compared to those of performance measures (i.e. the product-based assessment and divergent thinking tests). Finally, many of the self-report measures of creativity yield significantly skewed scores (e.g. CBI, CAQ and BICB), which calls for special statistical methods to manipulate this type of data (Silvia et al., 2012).

Product approach. The product-based assessment of creativity has been widely recommended for assessing creativity (Amabile, 1982; Baer, 1994c; Baer & McKool, 2009; Han, 2003; Horn & Salvendy, 2006; Kaufman et al., 2012). It has been argued that a comprehensive assessment of an individual's creativity cannot be attained without a measure of its product (Horn & Salvendy, 2006). The focus within this approach lies on rating individuals' creative products in different areas, such as writing, art, music, science or mathematics (O'Quin & Besemer, 1989). The Consensual Assessment Technique (CAT), suggested by Amabile (1982), has been frequently used in creativity studies that employed a product-based assessment (e.g. Amabile, 1982; Baer, 1994c; Baer, Kaufman, & Gentile, 2004; Christiaans, 2002; Kaufman et al., 2007; Tan et al., 2015). The CAT is not built upon any creativity theories; instead, its main idea is that the best evaluators of the quality of creative products are recognized experts in the relevant domain (Amabile, 1982; Hennessey, Amabile, & Mueller, 2011). Accordingly, the validity of the CAT is not subject to the acceptance or rejection of creativity theories whenever additional evidence is gathered (Baer, 1994c; Baer & McKool, 2009; Kaufman et al., 2007). As a result, it is easy for the CAT, unlike other creativity measures, to accommodate diverse perspectives on many of the controversial theoretical issues in creativity research.

The procedures of the CAT are similar to those in which real-life creativity is judged, and this is perhaps the reason why it is called the "gold standard" for creativity assessment (Baer & McKool, 2014). In the CAT, subjects are given tasks that require them to create something (a poem, a drawing, a collage, a story or a caption) (Kaufman et al., 2007). The task should be open-ended to allow more flexibility and novelty in responses,

lead to an observable product to be judged and not rely heavily on specific skills such as verbal facility or drawing ability (Amabile, 1982). Experts in the domain of interest are then asked to rate the level of creativity of each product on a Likert scale ranging from not at all creative to very creative (Barbot et al., 2011; Hennessey et al., 2011; Zhou & Shalley, 2003). Experts rate the products by comparing them against each other, rather than against an absolute standard (Amabile, 1982; Baer, 2012; Hennessey et al., 2011; Horn & Salvendy, 2006; Kaufman et al., 2011; Kaufman et al., 2012). They are informed to work independently, without any instructions or group discussions, and use their own experienced sense of what is creative in making their judgments (Baer, 2012, 2015; Baer & McKool, 2009; Hennessey et al., 2011; Kaufman & Baer, 2012; Kaufman et al., 2011). Once the ratings are obtained, the degree of agreement among raters is examined using inter-rater reliability scores. Finally, if the required reliability is achieved, the scores given by each expert to each product can be averaged to obtain the creativity score of a product (Amabile, 1982; Baer, 2015; Barbot et al., 2011; Hennessey et al., 2011).

Studies using the CAT for measuring creativity have reported high inter-rater reliabilities, generally ranging from .72 to .96 (e.g. Amabile, 1982; Baer, 1991, 1994b, 1994c; Brinkman, 1999). However, some situational variables seem to influence this reliability, such as the number of tasks being judged (Lee, Lee, & Youn, 2005; Kaufman et al., 2007) and the performance domain (Baer, 1991; Baer et al., 2004). In addition, the typically used methods (e.g. alpha and intraclass correlation coefficients) for assessing the inter-rater reliability of the CAT ratings only consider one source of measurement errors (i.e. judges) (Kaufman et al., 2007). These methods do not provide any information about the reliability estimates across various sources of measurement errors (e.g. task, domain and time), which limits the researcher's ability to draw inferences about the ideal measurement design to get an optimal level of reliability. But even leaving this aside, a high agreement among judges alone does not ensure appropriate conclusions about the quality of the measurement because it is only one criterion (Kaufman & Baer, 2012).

In addition, considerable concerns have been raised about the feasibility of using the CAT for measuring creativity. First, there is an ongoing debate about the selection of the most appropriate judges and the level of expertise needed to evaluate creative products (Galati, 2015; Hickey, 2001; Kaufman & Baer, 2012; Runco, McCarthy, & Svenson, 1993). It has been shown that the consensus among judges is affected by the expertise of these judges (Amabile, 1982; Christiaans, 2002; Hekkert & Van Wieringen, 1996; Hickey, 2001; Kaufman et al., 2008). Additionally, judges' differences in personality might influence creativity ratings. Tan et al. (2015) explored the effect of judges' personality, ac-

According to the big five model, on creativity ratings and reported that judges with high levels of agreeableness tended to be more lenient in their ratings. Also, cultural differences among judges might affect creativity ratings. Niu and Sternberg's (2001) investigation of the cultural influences on artistic creativity evaluation revealed that American judges were more severe and gave lower ratings than did Chinese judges. Additionally, judges might exhibit bias when evaluating their own products (intrapersonal evaluation) in comparison to evaluating products produced by others (interpersonal evaluation) (Runco & Smith, 1992; Runco et al., 1993). Runco and Smith's (1992) study indicated that judges were more accurate when rating the originality of their own ideas than when rating the originality of ideas produced by other people. However, they also found that judges were less accurate when rating the popularity of their own ideas than when rating the popularity of others' ideas. Furthermore, the discriminant power of judges' ratings of creative products in some domains is questionable. Lindauer (1991) asked judges of different ages and art background to compare mass-produced (cheap and less skillful art) and museum art. The results indicated that judges evaluated these two kinds of art similarly and failed to recognize the differences between them. Also, some studies have indicated that the agreement among experts in some domains is relatively low and no greater than agreement among non-experts or those with intermediate expertise (e.g. Christiaans, 2002; Hekkert & Van Wieringen, 1996; Lindauer & Long, 1986). A possible reason for the lack of consensus among experts is that although we can assume that experts target the same criteria in rating creativity, they may differ in their personal interpretation of these criteria and also in the relative weight that they allocate to each criterion (Lindauer & Long, 1986). These differences among experts may result in different subjective indicators and conflicting ordering of priorities in judging creativity. Moreover, economically speaking, gathering appropriate experts to evaluate creative products may be both expensive and time-consuming (Batey et al., 2010; Kaufman et al., 2007).

Second, the focus for measuring creativity using the CAT is entirely on the creative products, without any investigation of the cognitive processes that lead to the creation of these products (Amabile, 1982; Baer & McKool, 2009; Batey, 2012; Batey et al., 2010; Piffer, 2012; Runco, 2004). It is supposed that an individual generates such a product through the utilization of cognitive processes within a certain environment. Thus, the CAT cannot be regarded as a comprehensive measure of creativity as it reflects only one facet of creativity. Finally, scoring creativity using the CAT depends on judging subjects' products against each other within a particular sample, so the scores are only comparable within the sample of products being judged (Baer, 1994c; Baer & McKool, 2009). It is therefore not possible, under the CAT, to develop agreed-upon norms or criteria to com-

pare the products of other samples (Baer, 1994c; Baer & McKool, 2009; Fishkin & Johnson, 1998; Horn & Salvendy, 2006). Accordingly, an identical product may be considered creative in one sample, context or culture and less creative in another with a different group of products and judges. This may help to explain the low correlations and domain-specificity of creativity scores evaluated using the CAT in different domains (e.g. Baer, 1991, 1994a, 1994b, Han, 2003; Han & Marvin, 2002).

Press approach. The press approach concentrates on the analysis of the work environment or climate where creativity is realized (Batey, 2012; Hennessey & Amabile, 2010). The measurement of these environmental factors is based on the idea that there is an indirect relationship between these factors and creativity. Previous studies have revealed that environmental factors can significantly affect individuals' intrinsic motivation to engage in an activity, which in turn affects their creativity (see review of Hennessey & Amabile, 2010).

Much of the work in this area has been conducted by organizational behaviour researchers (see reviews of James et al., 1999; Mathisen & Einarsen, 2004; Shalley, Zhou & Oldham, 2004). They tend to study the effects of environmental factors on creativity at the individual (e.g. Amabile, Schatzel, Moneta, Kramer, & 2004; Forrester & Hui, 2007), organizational (e.g. Besançon & Lubart, 2008; Thomas & Berk, 1981) or cultural level (e.g. Leung, Maddux, Galinsky, & Chiu, 2008; Rudowicz, 2003). Numerous factors have been found to influence the development of creativity in learning environments. In a systematic review of studies relating to creative learning environments, Davies et al. (2013) determined key characteristics of school environments that enhance creativity, including an appropriate physical environment, availability of various materials, using outdoor environments, an exciting pedagogical environment, employing games-based approaches, flexible use of time and supportive relationships between teachers and students.

Commonly used instruments in this area include the Siegel Scale of Support of Innovation (SSSI; Siegel & Kaemmerer, 1978), the College and University Classroom Environment Inventory (CUCEI; Fraser, Treagust, & Dennis, 1986); the Work Environment Scale (WES; Moos, 1986); the Work Environment Inventory (WEI; Amabile & Gryskiewicz, 1989); KEYS: Assessing the Climate for Creativity (KEYS; Amabile, Conti, Coon, Lazenby, & Herron, 1996); the Team Climate Inventory (TCI; Anderson & West, 1998) and the Situational Outlook Questionnaire (SOQ; Isaksen, Lauer, & Ekvall, 1999). Generally, these instruments share some dimensions such as members' perceptions of organizational vision, resources, leadership, freedom, interactions, support for creativity and innovation and continuous development.

Instruments for measuring the work environment for creativity may focus on the degree to which this environment encourages or inhibits creativity. It may also focus on the strengths and weakness within different workgroups, evaluate environmental improvement attempts and take the corrective action necessary to ensure the support of creativity (Amabile et al., 1996). However, the existing efforts to measure creativity pay greater attention to other aspects of creativity (i.e. process, product and person) than to the climate in which creativity occurs (Craft, 2001; Hoelcher & Schubert, 2015). These efforts seem to consider only the internal or within-individual elements of creativity, excluding external elements such as the environmental factors (Amabile, 1983). As a result, much remains unknown about the effectiveness of this approach in measuring creativity and also about the psychometric properties of the instruments. Therefore, more studies are needed to evaluate the quality of this approach and its instruments (Amabile et al., 1996; Mathisen & Einarsen, 2004; Shalley et al., 2004). Another issue of concern with regard to this approach is the controversy about the meaning of “climate” (Anderson & West, 1998; Mathisen & Einarsen, 2004). Different definitions and dimensions have been proposed for this term (see, for example, Denison, 1996; Patterson et al., 2005). In this vein, Schneider and Reichers (1983) confirmed that the development of climate measures should be based on specific or definite referents (e.g. climate for quality, climate for development or climate for productivity etc.); otherwise, it would be surprising to find differences among groups or correlations with other variables through these measures. There also exists controversy about the level of climate that should be measured; is it better to measure the whole climate of an organization, or perhaps it is better to measure the climate at a team or workgroup level (Mathisen & Einarsen, 2004)? A final concern about this approach is that individuals may have different perceptions of climate as they conceptualize or understand situations differently according to their own personality, education or culture. Thus, the effect of the same climate on individuals’ creativity may be different due to their distinct internal representations of the climate, not to the climate per se.

CONCLUSIONS

The current review examined the literature on the measurement of creativity up to 2016. Our search strategy yielded 152 papers addressing the approaches and instruments used to measure creativity. Four major approaches and 18 instruments commonly used in these approaches were identified. These approaches are, in order of their frequency of use, the process, person, product and press approach. The findings of this review suggest that although the measurement of creativity has received much attention over the past several years, its methodology still has a long way to go in order to reach the desired

destination. Based on this review of the creativity literature, it seems that the field has long been troubled by a lack of consensus on the definition of creativity due to its multifaceted nature, which, in turn, has dispersed researchers' efforts to measure this construct. Consequently, numerous instruments to measure creativity can be found in literature. However, none of these instruments is without its limitations or can alone undertake the task of measuring the multifaceted construct of creativity. The common limitations of these instruments include the dissimilar and limited scope of measurement, conflicting evidence of validity, lack of appropriate norms and disregarding domain-relevant aspects of creativity. In addition, these instruments direct greater attention to within-individual aspects of creativity, ignoring the external context where creativity occurs.

The findings of our review add to the growing body of creativity research. The review offers a comprehensive contemporary overview of the existing literature on measuring creativity identified through scanning multiple sources of information. In this way, the insights gained from the review could be used to direct the measurement of creativity in research and practice. This review informs researchers about the latest developments in this research area and uncovers gaps and paths for subsequent investigations. It might also serve as a helpful guide for researchers and educators interested in measuring creativity for drawing contrasts between the advantages and weaknesses of the available instruments.

The findings of our review should be read in the light of some limitations. First, our search process included only studies written in English, which might have resulted in the exclusion of literature that could offer additional insights into this topic. We also cannot discount the possibility that relevant instruments for measuring creativity developed in other languages and/or used in other linguistic regions and cultures might have been missed. Second, like any review, our review is confined by the limitations of the primary studies included in our analysis. For instance, conceptual or methodological limitations of the selected studies might have distorted our conclusions. Finally, the scope of this review is limited to the quantitative approaches to measuring creativity. Qualitative measurement approaches of creativity (e.g. using interviews, case studies or ethnography) were left out.

According to this review, there are some suggestions that might help in addressing the limitations and unsolved issues related to the measurement of creativity. First, more than ever, it seems necessary to develop an agreed-upon conception of creativity that clearly delineates the nature of this construct and its components in view of its complexity. Following the development of this conception, the measurement of creativity has to be closely tied to this conception. Second, further investigations of the psychometric properties of the existing creativity instruments are needed to either support the psychometric

quality of these instruments or highlight the need for developing more reliable and valid instruments. Third, given the complexity of creativity, depending upon a single instrument to evaluate an individual's creativity might be insufficient or simply misleading. More research is needed on how instruments could work together to produce an individual's creativity profile that includes assessments of an individual's level on the different components of the creativity construct. In this regard, further work is required to expand upon the commonalities and differences among creativity instruments and to investigate whether the integration of these instruments could enhance their effectiveness and measurement outcomes. Finally, it would be beneficial to keep an eye on the variability of an individual's creativity profile across contexts, domains and age stages. Developing methodologies to investigate and understand the variability of creativity might widen our knowledge regarding this issue. This could lead to a dynamic conception of the construct of creativity rather than the popular static conception.

In conclusion, the measurement of creativity still appears to be an unsettled issue. The review of literature presented here, however, may offer valuable insights and provide starting points for concrete steps towards a better measurement of this construct.

REFERENCES

References marked with an asterisk were included in the analysis

- Abbey, A., & Dickson, J. W. (1983). R&D work climate and innovation in semiconductors. *Academy of Management Journal*, 26, 362-368.
- *Almeida, L. S., Prieto, M. D., Ferrando, M., Oliveira, E., & Ferrándiz, C. (2008). Torrance Test of Creative Thinking: The question of its construct validity. *Thinking Skills and Creativity*, 3, 53-58. doi: 10.1016/j.tsc.2008.03.003
- *Amabile, T. M. (1982). Social psychology of creativity: A Consensual Assessment Technique. *Journal of Personality and Social Psychology*, 43, 997-1013.
- Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, 45, 357-376.
- Amabile, T. M. (1988). A model of creativity and innovation in organizations. In B. M. Staw, & L. L. Cummings (Eds.), *Research in organizational behavior* (pp. 123-167). Greenwich, CT: JAI Press.
- Amabile, T. M. (1996). Evidence to support the componential model of creativity: Secondary analyses of three studies. *Creativity Research Journal*, 9, 385-389.
- *Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the work environment for creativity. *The Academy of Management Journal*, 39, 1154-1184.
- *Amabile, T. M., & Gryskiewicz, N. (1989). The creative environment scales: Work Environment Inventory. *Creativity Research Journal*, 2, 231-253.

- Amabile, T. M., Schatzel, E., Moneta, G. B., & Kramer, S. J. (2004). Leader behaviors and the work environment for creativity: Perceived leader support. *The Leadership Quarterly, 15*, 5-32.
- *Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: Development and validation of the team climate inventory. *Journal of Organizational Behavior, 19*, 235-258.
- *Bachlor, P. (1989). Maximum likelihood confirmatory factor-analytic investigation of factors within Guilford's Structure of Intellect model. *Journal of Applied Psychology, 74*, 797-804.
- *Bachlor, P., & Michael, W. B. (1991). Higher-order factors of creativity within Guilford's structure-of-intellect model: A re-analysis of a fifty-three variable data base. *Creativity Research Journal, 4*, 157-175.
- *Bachlor, P., Michael, W. B., & Kim, S. (1994). First-order and higher-order semantic and figural factors in Structure-of-Intellect Divergent Production measures. *Educational and Psychological Measurement, 54*, 608-619.
- *Baer, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal, 4*, 23-39.
- Baer, J. (1993). *Creativity and divergent thinking: A task-specific approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- *Baer, J. (1994a). Divergent thinking is not a general trait: A multi-domain training experiment. *Creativity Research Journal, 7*, 35-46.
- *Baer, J. (1994b). Generality of creativity across performance domains: A replication. *Perceptual and Motor Skills, 79*, 1217-1218.
- *Baer, J. (1994c). Performance assessments of creativity: Do they have long-term stability? *Roeper Review, 7*, 7-11.
- Baer, J. (1994d). Why you still shouldn't trust creativity tests. *Educational Leadership, 52*, 72-73.
- *Baer, J. (2011). How divergent thinking tests mislead us: Are the Torrance Tests still relevant in the 21st century? The division 10 debate. *Psychology of Aesthetics, Creativity, and the Arts, 5*, 309-313. doi: 10.1037/a0025210
- *Baer, J. (2012). Domain specificity and the limits of creativity theory. *The Journal of Creative Behavior, 46*, 16-29. doi: 10.1002/jocb.002
- *Baer, J. (2015). The importance of domain-specific expertise in creativity. *Roeper Review, 37*, 165-178. doi: 10.1080/02783193.2015.1047480
- *Baer, J. (2016). *Domain specificity of creativity*. San Diego, CA: Elsevier Academic Press.

- *Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the Consensual Assessment Technique to nonparallel creative products. *Creativity Research Journal*, 16, 113-17.
- *Baer, J., & McKool, S. S. (2009). Assessing creativity using the Consensual Assessment Technique. In C. S. Schreiner (Ed.), *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 65-77). Hershey, PA: IGI Global.
- *Baer, J., & McKool, S. S. (2014). The gold standard for assessing creativity. *International Journal of Quality Assurance in Engineering and Technology Education*, 3, 81-93. doi: 10.4018/ijqaete.2014010104
- *Barbot, B., Besançon, M., & Lubart, T. I. (2011). Assessing creativity in the classroom. *The Open Education Journal*, 4, 124-132. doi: 10.2174/1874920801104010058
- *Barbot, B., Besançon, M., & Lubart, T. I. (2015). Creative potential in educational settings: Its nature, measure, and nurture. *Education 3-13*, 43, 371-381. doi: 10.1080/03004279.2015.1020643
- *Barron, F., & Harrington, D. (1981). Creativity, intelligence, and personality. *Annual Review of Psychology*, 32, 439-476.
- *Batey, M. (2007). *A psychometric investigation of everyday creativity*. Unpublished Doctoral Thesis. London: University College.
- *Batey, M. (2012). The measurement of creativity: From definitional consensus to the introduction of a new heuristic framework. *Creativity Research Journal*, 24, 55-65. doi: 10.1080/10400419.2012.649181
- *Batey, M., Chamorro-Premuzic, T., & Furnham, A. (2010). Individual differences in ideational behavior: Can the big five and psychometric intelligence predict creativity scores? *Creativity Research Journal*, 22, 90-97. doi: 10.1080/10400410903579627
- *Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, 132, 355-429.
- *Belcher, T. L., Rubovits, J. J., & Di Meo, P. A. (1981). Measurement of creativity: A factor analytic study. *Psychological Reports*, 48, 819-825.
- Besançon, M., & Lubart, T. I. (2008). Differences in the development of creative competencies in children schooled in diverse learning environments. *Learning and Individual Differences*, 18, 381-389. doi: 10.1016/j.lindif.2007.11.009
- Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist*, 58, 1019-1027.
- Blamiresa, M., & Peterson, A. (2014). Can creativity be assessed? Towards an evidence-informed framework for assessing and planning progress in creativity. *Cambridge Journal of Education*, 44, 147-162. doi: 10.1080/0305764X.2013.860081

- *Brinkman, D. (1999). Problem finding, creativity style and the musical compositions of high school students. *The Journal of Creative Behavior*, 33, 62-68.
- *Brougher, S. J., & Rantanen, E. M. (2009). Creativity and design: Creativity's new definition and its relationship to design. *Proceedings of the Human Factors and Ergonomics Society*, 1, 605-609. doi: 10.1177/154193120905301005
- *Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal*, 17, 37-50.
- Christiaans, H. H. C. M. (2002). Creativity as a design criterion. *Creativity Research Journal*, 14, 41-54.
- *Clapham, M. M. (1998). Structure of the figural forms A and B of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*, 58, 275-283.
- *Clapham, M. M. (2004). The convergent validity of the Torrance Tests of Creative Thinking and creativity interest inventories. *Educational and Psychological Measurement*, 64, 828-841.
- Colangelo, N., Kerr, B., Hallowell, K., Huesman, R., & Gaeth, J. (1992). The Iowa inventiveness inventory: Toward a measure of mechanical inventiveness. *Creativity Research Journal*, 5, 157-163.
- *Couger, J. D., Higgins, L. F., & McIntyre, S. C. (1993). (Un)structured creativity in information systems organizations. *MIS Quarterly*, 17, 375-397.
- Craft, A. (2001). *An analysis of research and literature on creativity in education*. Report prepared for the qualifications and curriculum authority. Available at http://www.creativetallis.com/uploads/2/2/8/7/2287089/creativity_in_education_report.pdf
- *Cramond, B., Matthews-Morgan, J., Bandalos, D., & Zuo, L. (2005). A report on the 40-year follow-up of the Torrance Tests of Creative Thinking: Alive and well in the new millennium. *Gifted Child Quarterly*, 49, 283-291.
- *Cropley, A. J. (1972). A five-year longitudinal study of DT tests. *Developmental Psychology*, 6, 119-124.
- *Cropley, A. J. (2000). Defining and measuring creativity: Are creativity tests worth using? *Roeper Review*, 23, 72-79.
- *Cropley, A. J., & Maslany, G. W. (1969). Reliability and factorial validity of the Wallach-Kogan Creativity Tests. *British Journal of Psychology*, 60, 395-398.
- *Davidovitch, N., & Milgram, R. M. (2006). Creative thinking as a predictor of teacher effectiveness in higher education. *Creativity Research Journal*, 18, 385-390.

- Davies, D., Jindal-Snape, D., Collier, C., Digby, R., Hay, P., & Howe, A. (2013). Creative learning environments in education-A systematic literature review. *Thinking Skills and Creativity, 8*, 80-91. doi: 10.1016/j.tsc.2012.07.004
- *Davis, G. A. (1975). In frumious pursuit of the creative person. *The Journal of Creative Behavior, 9*, 75-87.
- *Davis, G. A., & Subkoviak, M. J. (1975). Multidimensional analysis of a personality-based test of creative potential. *Journal of Educational Measurement, 12*, 37-43.
- Denison, D. R. (1996). What is the difference between organizational culture and organizational climate? A native's point of view on a decade of paradigm wars. *The Academy of Management Review, 21*, 619-654.
- *Diakidoy, I. N., & Spanoudis, G. (2002). Domain specificity in creativity testing: A comparison of performance on a general divergent-thinking test and a parallel, content-specific test. *The Journal of Creative Behavior, 36*, 41-61.
- *Dixon, J. (1979). Quality versus quantity: The need to control for the fluency factor in originality scores from the Torrance Tests. *Journal for the Education of the Gifted, 2*, 70-79.
- Ericsson, K. A. (1999). Creative expertise as superior reproducible performance: Innovative and flexible aspects of expert performance. *Psychological Inquiry, 10*, 329-333.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Bulletin, 2*, 290-309.
- Feist, G. J. (1999). The influence of personality on artistic and scientific creativity. In R. J. Sternberg (Ed.), *Handbook of creativity*. New York: Cambridge University Press.
- Finke, R. A., Ward, T. M., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.
- *Fishkin, A. S., & Johnson, A. S. (1998). Who is creative? Identifying children's creative abilities. *Roeper Review, 21*(1), 40-46. doi: 10.1080/02783199809553925
- *Fleenor, J. W., & Taylor, S. (1994). Construct validity of three self-report measures of creativity. *Educational and Psychological Measurement, 54*, 464-470.
- Forrester, V., & Hui, A. (2007). Creativity in the Honk Kong classroom: What is the contextual practice? *Thinking Skills and Creativity, 2*, 30-38.
- *Fraser, B. J., Treagust, D. F., & Dennis, N. (1986). Development of an instrument for assessing classroom psychosocial environment at universities and colleges. *Studies in Higher Education, 11*, 43-54.

- *Furnham, A., Batey, M., Anand, K., & Manfield, J. (2008). Personality, hypomania, intelligence and creativity. *Personality and Individual Differences, 44*, 1060-1069. doi: 10.1002/j.2162-6057.2008.tb01293.x
- Gajda, A., Karwowski, M., & Beghetto, R. A. (2016). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology*. Advance online publication. doi. 10.1037/edu0000133
- *Galati, F. (2015). Complexity of judgment: What makes possible the convergence of expert and nonexpert ratings in assessing creativity. *Creativity Research Journal, 27*, 24-30. doi: 10.1080/10400419.2015.992667
- Gardner, H. (1993). *Creating minds*. New York: Basic Books.
- Geis, G.T. (1988). Making Companies Creative. In R. L. Kuhn (ed.), *Handbook for creative and innovative managers* (pp. 25-33). New York: McGraw-Hill.
- *Getzels, J. W., & Csikszentmihalyi, M. (1976) *The creative vision: A longitudinal study of problem finding in Art*. New York: Wiley.
- *Gluck, J., Ernst, R., & Unger, F. (2002). How creatives define creativity: Definitions reflect different types of creativity. *Creativity Research Journal, 14*, 55-67.
- *Goldsmith, R. E., & Matherly, T. A. (1988). Creativity and self-esteem: A multiple operationalization validity study. *The Journal of Psychology, 122*, 47-56.
- *González, M. A., Campos, A., & Pérez, M. J. (1997). Mental imagery and creative thinking. *The Journal of Psychology, 131*, 357-364.
- *Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology, 37*, 1398-1405.
- Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444-454.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Guilford, J. P. (1975). Creativity: A quarter century of progress. In I. A. Taylor, & J. W. Getzels (Eds.), *Perspectives in creativity* (pp. 37-59). Chicago: Aldine.
- *Han, K. (2003). Domain-specificity of creativity in young children: How quantitative and qualitative data support it. *The Journal of Creative Behavior, 37*, 117-142.
- *Han, K., & Marvin, C. (2002). Multiple creativities? Investigating domain-specificity of creativity in young children. *Gifted Child Quarterly, 46*, 98-109.
- Harrington, D. (1975). Effects of explicit instructions to "be creative" on the psychological meaning of divergent thinking test scores. *Journal of Personality, 43*, 434-454.
- Harrington, D. (1990). The ecology of human creativity: A psychological perspective. In M. Runco, & R. Albert (Eds.), *Theories of creativity* (pp. 143-169). Newbury Park, CA: SAGE.

- *Heausler, N. L., & Thompson, B. (1988). Structure of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement, 48*, 463-468.
- Hekkert, P., & Van Wieringen, P. C. W. (1996). Beauty in the eye of expert and nonexpert beholders: A study in the appraisal of art. *American Journal of Psychology, 109*, 389-407.
- *Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology, 61*, 569-598. doi: 10.1146/annurev.psych.093008.100416
- *Hennessey, B. A., Amabile, T. M., & Mueller, J. S. (2011). Consensual Assessment. In M. A. Runco, & S. R. Pritzker (Eds.) *Encyclopedia of creativity*, vol. 1 (2nd ed.) (pp. 253-260). San Diego: Academic Press.
- *Hickey, M. (2001). An application of Amabile's Consensual Assessment Technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education, 49*, 234-244.
- *Hocevar, D. (1979a). A comparison of statistical infrequency and subjective judgment as criteria in the measurement of originality. *Journal of Personality Assessment, 43*, 297-299.
- *Hocevar, D. (1979b). Ideational fluency as a confounding factor in the measurement of originality. *Journal of Educational Psychology, 71*, 191-196.
- *Hocevar, D. (1979c). *The development of the Creative Behavior Inventory (CBI)*. Paper presented at the 1979 conference of the Rocky Mountain Psychological Association, Las Vegas, Nevada (ERIC Document Reproduction Service No. ED170350).
- *Hocevar, D. (1979d). The unidimensional nature of creative thinking in fifth grade children. *Child Study Journal, 9*, 273-278.
- *Hocevar, D. (1980). Intelligence, divergent thinking, and creativity. *Intelligence, 4*, 25- 40.
- *Hocevar, D. (1981). Measurement of creativity: Review and critique. *Journal of Personality Assessment, 45*, 450-464.
- *Hocevar, D., & Bachelor, P. (1989). A taxonomy and critique of measurements 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 Press.
- *Hocevar, D., & Michael, W. B. (1979). The effects of scoring formulas on the discriminant validity of tests of divergent thinking. *Educational and Psychological Measurement, 39*, 917-921.
- Hoelcher, M., & Schubert, M. (2015). Potential and problems of existing creativity and innovation indices. *Creativity Research Journal, 27*, 1-15. doi: 10.1080/10400419.2015.992656

- Holland, J. L., & Nichols, R. (1964). Prediction of academic and extracurricular achievement in college. *Journal of Educational Psychology, 55*, 55-65.
- *Horn, D., & Salvendy, G. (2006). Consumer-based assessment of product creativity: A review and reappraisal. *Human Factors and Ergonomics in Manufacturing, 16*, 155-175.
- *Isaksen, S. G., Lauer, K. J., & Ekvall, G. (1999). Situational Outlook Questionnaire: A measure of the climate for creativity and change. *Psychological Reports, 85*, 665-674.
- James, K., & Asmus, C., (2000-2001). Personality, cognitive skills, and creativity in different life domains. *Creativity Research Journal, 13*, 149-159.
- James, K., Clark, K., & Cropanzano, R. (1999). Positive and negative creativity in groups, institutions, and organizations: A model and theoretical extension. *Creativity Research Journal, 12*, 211-226.
- *Jauk, E., Benedek, M., & Neubauer, A. C. (2014). The road to creative achievement: A latent variable model of ability and personality predictors. *European Journal of Personality, 28*, 95-105. doi: 10.1002/per.1941
- *Kaufman, J. C. (2012). Counting the Muses: Development of the Kaufman Domains of Creativity Scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts, 6*, 298-308. doi: 10.1037/a0029751
- *Kaufman, J. C., & Baer, J. (2004). Hawking's Haiku, Madonna's math: Why it is hard to be creative in every room of the house. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 3-19). Washington, DC: American Psychological Association.
- *Kaufman, J. C., & Baer, J. (2012). Beyond new and appropriate: Who decides what is creative? *Creativity Research Journal, 24*, 83-91. doi: 10.1080/10400419.2012.649237
- *Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the Consensual Assessment Technique. *Creativity Research Journal, 20*, 171-178. doi: 10.1080/10400410802059929
- *Kaufman, J. C., Evans, M. L., & Baer, J. (2010). The American Idol Effect: Are students good judges of their creativity across domains? *Empirical Studies of the Arts, 28*, 3-17. doi: 10.2190/EM.28.1.b
- *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.
- *Kaufman, J. C., Plucker, J. A., & Russell, C. M. (2012). Identifying and assessing creativity as a component of giftedness. *Journal of Psychoeducational Assessment, 30*, 60-73. doi: 10.1177/0734282911428196

- Kaufman, J. C., & Sternberg, R. J. (2007). Resource review: Creativity. *Change*, 39, 55-58.
- *Khatena, J., & Torrance, E. P. (1973). *Thinking creatively with sounds and words: Technical Manual (Research Ed.)*. Lexington, MA: Personnel Press.
- *Khattab, A., Michael, W. B., & Hocevar, D. (1982). The construct validity of higher order Structure-of-Intellect abilities in a battery of tests emphasizing the product of transformations: A confirmatory maximum likelihood factor analysis. *Educational and Psychological Measurement*, 42, 1089-1105.
- *Kim, K. H. (2006a). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, 18, 3-14.
- *Kim, K. H. (2006b). Is creativity unidimensional or multidimensional? Analyses of the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 18, 251-259.
- Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *The Journal of Creative Behavior*, 42, 106-130. doi: 10.1002/j.2162-6057.2008.tb01290.x
- *Kim, K. H., Cramond, B., & Bandalos, D. L. (2006). The latent structure and measurement invariance of scores on the Torrance Tests of Creative Thinking-figural. *Educational and Psychological Measurement*, 66, 459-477.
- *King, L. A., Walker, L. M., & Broyles, S. J. (1996). Creativity and the five-factor model. *Journal of Research in Personality*, 30, 189-203.
- *Kogan, N., & Pankove, E. (1974). Long-term predictive validity of divergent-thinking tests: Some negative evidence. *Journal of Educational Psychology*, 66, 802-810.
- *Krumm, G., Aranguren, M., Filippetti, V. A., & Lemos, V. (2014). Factor structure of the Torrance Tests of Creative Thinking verbal form B in a Spanish-speaking population. *The Journal of Creative Behavior*, 50, 150-164. doi: 10.1002/jocb.76
- *Krumm, G., Filippetti, V. A., Lemos, V., Koval, J., & Balabanian, C. (2016). Construct validity and factorial invariance across sex of the Torrance Test of Creative Thinking – figural Form A in Spanish-speaking children. *Thinking Skills and Creativity*, 22, 180-189. doi: 10.1016/j.tsc.2016.10.003
- *Krumm, G., Lemos, V., & Filippetti, V. A. (2014). Factor structure of the Torrance Tests of Creative Thinking figural form B in Spanish-speaking children: Measurement invariance across gender. *Creativity Research Journal*, 26, 72-81. doi: 10.1080/10400419.2013.843908
- *Lee, S., Lee, J., & Youn, C.-Y. (2005). A variation of CAT for measuring creativity in business products. *Korean Journal of Thinking and Problem Solving*, 15, 143-153.

- *Lemons, G. (2011). Diverse perspectives of creativity testing: Controversial issues when used for inclusion into gifted programs. *Journal for the Education of the Gifted*, 34, 742-772. doi: 10.1177/0162353211417221
- Leung, A. K.-Y., Maddux, W. W., Galinsky, A. D., & Chiu, C.-Y. (2008). Multicultural experience enhances creativity: The when and how. *American Psychologist*, 63, 169-181. doi: 10.1037/0003-066X.63.3.169
- Lindauer, M. S. (1991). Comparisons between museum and mass-produced art. *Empirical Studies of the Arts*, 9, 11-22.
- Lindauer, M. S., & Long, D. A. (1986). The criteria used to judge art: Marketplace and academic comparisons. *Empirical Studies of the Arts*, 4, 163-174.
- Lubart, T. I. (2001). Models of the creative process: Past, present and future. *Creativity Research Journal*, 13, 295-308.
- Lubart, T. I., & Guignard, J. (2004). The generality-specificity of creativity: A multivariate approach. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 43-56). Washington, DC: American Psychological Association.
- *Lubart, T. I., Zenasni, F., & Barbot, B. (2013). Creative potential and its measurement. *International Journal for Talent Development and Creativity*, 1, 41-51.
- *Mace, D. E., Michael, W. B., & Hocevar, D. (1985). Validity of higher-order ability constructs in Structure-Of-Intellect tests all involving semantic content and operations of cognition or evaluation: A confirmatory maximum likelihood factor analysis. *Educational and Psychological Measurement*, 45, 353-359.
- *Mathisen, G. E., & Einarsen, S. (2004). A review of instruments assessing creative and innovative environments within organizations. *Creativity Research Journal*, 16, 119-140.
- Mayer, R. E. (1999). Fifty years of Creativity Research. In R. J. Sternberg (Ed.) *Handbook of Creativity* (pp. 449-460). New York: Cambridge University Press.
- McLaren, R. B. (1993). The dark side of creativity. *Creativity Research Journal*, 6, 137-144.
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, 69, 220-232.
- *Michael, W. B., & Bachelor, P. (1990). Higher-order structure-of-intellect creativity factors in divergent production tests: A re-analysis of a Guilford data base. *Creativity Research Journal*, 3, 58-74.
- *Milgram, R. M., & Hong, E. (1993). Creative thinking and creative performance in adolescents as predictors of creative attainments in adults: A follow-up study after 18 years. *Roeper Review*, 15, 135-139.

- *Miller, A. L. (2014). A self-report measure of cognitive processes associated with creativity. *Creativity Research Journal*, 26, 203-218. doi: 10.1080/10400419.2014.901088
- *Moos, R. (1986). *Work Environment Scale: Manual*. Palo Alto, CA: Consulting Psychologists Press.
- *Mouchiroud, C., & Lubart, T. (2001). Children's original thinking: An empirical examination of alternative measures derived from divergent thinking tasks. *The Journal of Genetic Psychology*, 162, 382-401.
- *Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creativity Research Journal*, 15, 107-120.
- Ng, T. W. H., & Feldman, D. C. (2012). A comparison of self-ratings and non-self-report measures of employee creativity. *Human Relations*, 65, 1021-1047. doi: 10.1177/0018726712446015
- *Niu, W., & Sternberg, R. J. (2001). Cultural influences on artistic creativity and its evaluation. *International Journal of Psychology*, 36, 225-241.
- Okuda, S. M., Runco, M. A., & Berger, D. E. (1991). Creativity and the finding and solving of real-world problems. *Journal of Psychoeducational Assessment*, 9, 45-53.
- *O'Quin, K., & Besemer, S. P. (1989). The development, reliability, and validity of the revised creative product semantic scale. *Creativity Research Journal*, 2, 267-278.
- *Palmiero, M., Nori, R., Aloisi, V., Ferrara, M., & Piccardi, L. (2015). Domain-specificity of creativity: A study on the relationship between visual creativity and visual mental imagery. *Frontiers in Psychology*, 6, 1-8. doi: 10.3389/fpsyg.2015.01870
- Paramithaa, A., & Indarti, N. (2014). Impact of the environment support on creativity: Assessing the mediating role of intrinsic motivation. *Procedia-Social and Behavioral Sciences*, 115, 102-114. doi: 10.1016/j.sbspro.2014.02.419
- *Park, N. K., Chun, M. Y., & Lee, J. (2016). Revisiting individual creativity assessment: Triangulation in subjective and objective assessment methods. *Creativity Research Journal*, 28, 1-10. doi: 10.1080/10400419.2016.1125259
- *Patterson, M. G., West, M. A., Shackleton, V. J., Dawson, J. F., Lawthom, R., Maitlis, S., Robinson, D. L., & Wallace, A. M. (2005). Validating the organizational climate measure: Links to managerial practices, productivity and innovation. *Journal of Organizational Behavior*, 26, 379-408.
- *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
- *Plucker, J. A. (1999a). Is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal data. *Creativity Research Journal*, 12, 103-114.

- *Plucker, J. A. (1999b). Reanalyses of student responses to creativity checklists: Evidence of content generality. *The Journal of Creative Behavior*, 33, 126-137.
- *Plucker, J. A., & Makel, M. C. (2010). Assessment of creativity. In J. Kaufman, & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity*. New York: Cambridge University Press.
- *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. *The Journal of Creative Behavior*, 45, 1-22. doi: 10.1002/j.2162-6057.2011.tb01081.x
- *Plucker, J. A., & Renzulli, J. S. (1999). Psychometric approaches to the study of creativity. In R. J. Sternberg (Ed.). *Handbook of human creativity* (pp. 35-60). New York: Cambridge University Press.
- *Plucker, J. A., & Runco, M. A. (1998). The death of creativity measurement directions in creativity assessment has been greatly exaggerated: Current issues, recent advances, and future. *Roepers Review*, 21, 36-39.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 879-903.
- *Priest, T. (2006). Self-evaluation, creativity, and musical achievement. *Psychology of Music*, 34, 47-61.
- 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* (online publication). doi: 10.1037/a0038055
- *Raudsepp, E. (1981). *How creative are you?* New York: Perigee.
- *Reiter-Palmon, R., Robinson-Morrall, E. J., Kaufman, J. C., & Santo, J. B. (2012). Evaluation of self-perceptions of creativity: Is it a useful criterion? *Creativity Research Journal*, 24, 107-114. doi: 10.1080/10400419.2012.676980
- *Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan*, 42, 109-135.
- Rudowicz, E. (2003). Creativity and culture: A two way interaction. *Scandinavian Journal of Educational Research*, 47, 273-290.
- *Runco, M. A. (1984). Teachers' judgments of creativity and social validation of divergent thinking tests. *Perceptual and Motor Skills*, 59, 711-717.
- *Runco, M. A. (1986). Divergent thinking and creative performance in gifted and nongifted children. *Educational and Psychological Measurement*, 46, 375-384.
- *Runco, M. A. (1987). The generality of creative performance in gifted and nongifted children. *Gifted Child Quarterly*, 31, 121-125.

- *Runco, M. A. (1993). Divergent thinking, creativity, and giftedness. *Gifted Child Quarterly*, 37, 16-22.
- *Runco, M. A. (2004). Creativity. *Annual Review of Psychology*, 55, 657-687.
- *Runco, M. A. (2007). *Creativity. Theories and themes: Research, development and practice*. Amsterdam: Elsevier.
- *Runco, M. A. (2008). Commentary: Divergent thinking is not synonymous with creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 2, 93-96. doi: 10.1037/1931-3896.2.2.93
- *Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, 24, 66-75. doi: 10.1080/10400419.2012.652929
- Runco, M. A., & Bahleda, M. D. (1986). Implicit theories of artistic, scientific, and everyday creativity. *The Journal of Creative Behavior*, 20, 93-98.
- Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24, 92-96. doi: 10.1080/10400419.2012.650092
- Runco, M. A., & Johnson, D. J. (1993). Parents' and teachers' implicit theories of children's creativity. *Child Study Journal*, 23, 91-113.
- *Runco, M. A., McCarthy, K. A., & Svenson, E. (1993). Judgments of the creativity of artwork from students and professional artists. *The Journal of Psychology*, 128, 23-31.
- *Runco, M. A., Millar, G., Acar, S., & Cramond, B. (2010). Torrance Tests of Creative Thinking as predictors of personal and public achievement: A fifty year follow up. *Creativity Research Journal*, 22, 361-368. doi: 10.1080/10400419.2010.523393
- *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.
- *Runco, M. A., & Okuda, S. M. (1988). Problem discovery, divergent thinking, and the creative process. *Journal of Youth and Adolescence*, 17, 211-220.
- *Runco, M. A., Okuda, S. M., & Thurston, B. J. (1987). The psychometric properties of four systems for scoring divergent thinking tests. *Journal of Psychoeducational Assessment*, 5, 149-156.
- *Runco, M. A., & Smith, W. R. (1992). Interpersonal and intrapersonal evaluations of creative ideas. *Personality and Individual Differences*, 13, 295-302.
- Schneider, B., & Reichers, A. E. (1983). On the etiology of climates. *Personnel Psychology*, 36, 19-39.
- *Shalley, C. E., Zhou, J., & Oldham, G. R. (2004). The effects of personal and contextual characteristics on creativity: Where should we go from here? *Journal of Management*, 30, 933-958.

- *Siegel, S. M., & Kaemmerer, W. F. (1978). Measuring the perceived support for innovation in organizations. *Journal of Applied Psychology*, 63, 553-562.
- Silvia, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychological Review* (online publication). 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., Wigert, B., Reiter-Palmon, R., & Kaufman, J. C. (2012). Assessing creativity with self-report scales: A review and empirical evaluation. *Psychology of Aesthetics, Creativity, and the Arts*, 6, 19-34. doi: 10.1037/a0024071
- *Skager, R. W., Klein, S. P., & Schultz, C. B. (1967). The prediction of academic and artistic achievement at a school of design. *Journal of Educational Measurement*, 4, 105-117.
- Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, 49, 607-627.
- Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, 41, 37-55.
- Sternberg, R. J., & Lubart T. I. (1999). The concept of creativity: Prospects and paradigms. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 3-15). New York: Cambridge University Press.
- Thomas, N. G., & Berk, L. E. (1981). Effects of school environments on the development of young children's creativity. *Child Development*, 52, 1153-1162.
- *Thompson, G., & Lordan, M. (1999). A review of creativity principles applied to engineering design. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 213, 17-31.
- *Torrance, E. P. (1966). *The Torrance Tests of Creative Thinking: Norms-technical manual*. Lexington, MA: Personal Press.
- *Torrance, E. P. (1969). Prediction of adult creative achievement among high school seniors. *Gifted Child Quarterly*, 13, 223-229.
- *Torrance, E. P. (1972a). Career patterns and peak creative achievements of creative high school students 12 years later. *Gifted Child Quarterly*, 16, 75-88.
- *Torrance, E. P. (1972b). Predictive validity of the Torrance Tests of Creative Thinking. *The Journal of Creative Behavior*, 6, 236-252.
- *Torrance, E. P. (1977). *Creativity in the classroom: What research says to the teacher*. Washington, DC: National Education Association.
- *Torrance, E. P. (1980). Growing up creatively gifted: The 22-year longitudinal study. *The Creative Child and Adult Quarterly*, 5, 148-158.

- *Torrance, E. P. (1981a). Empirical validation of criterion referenced indicators of creative ability through a longitudinal study. *The Creative Child and Adult Quarterly*, 6, 136-140.
- *Torrance, E. P. (1981b). Predicting the creativity of elementary school children (1958-80) and the teacher who made a difference. *Gifted Child Quarterly*, 25, 55-62.
- *Torrance, E. P. (2008). *The Torrance Tests of Creative Thinking-norms-technical manual-figural (streamlined) forms A and B*. Bensenville, IL: Scholastic Testing Service.
- *Torrance, E. P., & Haensly, P. A. (2003). Assessment of creativity in children and adolescents. In C. R. Reynolds, & R. W. Kamphaus (Eds.) *Handbook of psychological and educational assessment of children: Intelligence, aptitude and achievement* (2nd ed.) (pp. 584-607). New York: The Guildford Press.
- Treffinger, D. J. (1996). *Creativity, creative thinking, and critical thinking: In search of definitions*. Sarasota, FL: Center for Creative Learning.
- *Treffinger, D. J., Renzulli, J., & Feldhusen, J. (1971). Problems in the assessment of creative thinking. *The Journal of Creative Behavior*, 5, 104-112.
- *Treffinger, D. J., Young, G. C., Selby, E. C., & Shepardson, C. (2002). *Assessing creativity: A guide for educators*. Sarasota: Center for Creative Learning.
- *Ulosevich, S. N., Michael, W. B., & Bachelor, P. (1991). Higher-order factors in Structure-of-Intellect (SOI) aptitude tests hypothesized to portray constructs of military leadership: A re-analysis of an SOI data base. *Educational and Psychological Measurement*, 51, 15-37.
- Ursyn, A. (2014). *Perceptions of Knowledge Visualization: Explaining Concepts through Meaningful Images*. Hershey, PA: IGI Global.
- Vaughn, V. L., Feldhusen, J. F., & Asher, J. W. (1991). Meta-analyses and review of research on pull-out programs in gifted education. *Gifted Child Quarterly*, 35, 92-98.
- *Vincent, P. H., Decker, B. P., & Mumford, M. D. (2002). Divergent thinking, intelligence, and expertise: A test of alternative models. *Creativity Research Journal*, 14, 163-178.
- *Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity intelligence distinction*. New York: Holt, Rinehart & Winston.
- *Wechsler, S. (2006). Validity of the Torrance Tests of Creative Thinking to the Brazilian culture. *Creativity Research Journal*, 18, 15-25.
- Weisberg, R. W. (2006). *Creativity: understanding innovation in problem solving, science, invention, and the arts*. Hoboken, N.J.: John Wiley & Sons.
- Westby, E. L., & Dawson, V. L. (1995). Creativity: Asset or burden in the classroom? *Creativity Research Journal*, 8, 1-10.
- *Williams, F. (1980). *Creativity assessment packet*. Buffalo, NY: DOK.

- Zeng, L., Proctor, R. W., & Salvendy, G. (2009). Fostering creativity in service development: Facilitating service innovation by the creative cognition approach. *Service Science*, 1, 142-153. doi: 10.1287/serv.1.3.142
- *Zeng, L., Proctor, R. W., & Salvendy, G. (2011). Can traditional divergent thinking tests be trusted in measuring and predicting real-world creativity? *Creativity Research Journal*, 23, 24-37. doi: 10.1080/10400419.2011.545713
- *Zhou, J., & Shalley, C. E. (2003). Research on employee creativity: A critical review and directions for future research. *Research in Personnel and Human Resources Management*, 22, 165-217.

Corresponding author at: Sameh Said-Metwaly, Faculty of Psychology and Educational Sciences, KU Leuven (University of Leuven), 8500 Kortrijk, Belgium.
E-mail: sameh.metwaly@kuleuven.be

