ICPE 2017
International Conference on Psychology and Education

CREATIVE STORYTELLING IN CHILDHOOD IS RELATED TO EXAM PERFORMANCE AT AGE 16

Teemu Toivainen (a), Margherita Malanchini (b), Bonamy R. Oliver (c),
Yulia Kovas (d)*
*Corresponding author

(a) Department of Psychology, Goldsmiths, University of London, UK
(b) Department of Psychology, Population Research Center, University of Texas at Austin, U.S. & Social, Genetic and Developmental Psychiatry Centre, King's College, London, UK
(c) Department of Psychology, Goldsmiths, University of London, UK
(d) Department of Psychology, Goldsmiths, University of London, UK; Department of Psychology, Goldsmiths, University of London, UK; Tomsk State University, Tomsk, Russia, e-mail: y.kovac@gold.ac.uk

Abstract

Creativity is only partly recognised in education. A recent meta-analysis estimated a correlation of $r = 0.22$ between creativity and educational achievement across many international student samples of all educational levels. In the meta-analysis, creativity was measured with a variety of measures, including divergent thinking and remote association tasks. The differences in the measures influenced the strength of the relationship between creativity and educational achievement. More research is needed to establish reliable measures of creativity, especially in primary school children, whose creativity remains poorly evaluated. The present study measured creativity in written stories in children at age 9 using the Consensual Assessment Technique (CAT). The study employed a longitudinal design, using CAT creativity scores as a predictor of educational achievement at age 16. Each of the stories from 59 children were coded by 6 different judges for 10 dimensions, including creativity. The inter-rater reliabilities between the judges for the 10 dimensions were high ($\alpha = .76 - .95$). Among the dimensions, a factor analysis revealed two factors: Creative Expressiveness and Logic. The Creative Expressiveness factor explained an additional 7 % of variance in English grades, but not in Maths, beyond intelligence, previous achievement and personality traits associated with creativity. Overall, the study showed that CAT is a robust and reliable measure to detect verbal creativity in childhood. The results also suggest that early creativity predicts later academic achievement, calling for more attention to early creativity assessment and development.

1. Introduction

Creativity has positive associations in contemporary societies. For example, creativity is essential for new ideas and innovations in business, and it is regarded as a central facet in art and design. In recent decades, the role of creativity has also been emphasised in the educational discourse (Craft, 2003). For example, recognition of creativity is listed as one of the aims of the National Curriculum in the UK (Department of Education, UK, 1994).

Creative idea production is recognised as being an outcome of both divergent and convergent thinking (e.g. Cropley, 2006). However, the education system in the UK emphasises convergent thinking skills (Wilson, 2014). Convergent thinking is characterised by having one correct solution to a clearly defined problem (Guilford, 1957). Most cognitive ability tests, as well as many exams in primary education, measure convergent thinking skills (Chamorro-Premuzic & Reichenbacher, 2008; Cropley, 1967). In contrast, divergent thinking tests have dominated the field of creativity assessment for decades (Runco & Acar, 2012). Divergent thinking involves producing multiple answers or alternative answers from available information, as well as making unexpected combinations, remote associations and transforming information (Cropley, 2006). Divergent thinking is associated with exploratory learning style (Dirkes, 1978). For example, Montessori schools emphasise creative learning through activities based on imagination (Besançon & Lubart, 2008). Research has also highlighted the teachers' role as facilitators of creative learning (Jeffrey, 2006). However, more research is needed in order to establish the ideal circumstances in education for creativity to flourish.

The assessment or evaluation of creativity in education is not easy, since there are many different ways to define creativity (Plucker, Beghetto & Dow, 2004). This problem was demonstrated in a recent meta-analysis on creativity and educational achievement. The meta-analysis concluded that the differences between studies in which creativity measures were used were reflected in differences across the studies in the strength of associations between creativity and educational achievement (Gajda, Karwowski & Beghetto, 2016). In other words, different creativity measures do not correlate highly among themselves, tapping into largely different aspects of creativity. The combined, overall effect size for all different creativity measures and educational achievement in the meta-analysis was r = .2 (Gajda et al., 2016).

The creativity measures used in the meta-analysis included self-evaluations of one's own creativity, frequency of taking part to creative activities, divergent thinking and insight tasks (Gajda et al., 2016). Insight tasks present participants with unusual problems that require an alternative, new way of addressing a problem. Arriving at the solution of an insight task is associated with a sudden and clear solution through insight, the 'A-ha' or 'Eureka' moment (Bowden et al., 2005). Insight tasks correlate poorly with other creativity measures (e.g. divergent thinking and behavioural measures), and they are conceptually very similar to many convergent thinking tasks (Beaty, Nusbaum, & Silvia, 2014). The most commonly used divergent thinking task is the Alternative Uses Task, such as the Torrance tests of Creative Thinking (Torrance, Ball, & Safter, 2003). These tasks require participants to come up with alternative uses to a shown object, such as brick or newspaper. These tasks typically require external, subjective evaluations to score the answers based on their creativity. The Torrance Tests may measure divergent thinking ability in certain specific domains, but they should not be interpreted as measures of

creativity in general (Baer, 2011). As the score is based on frequency of the answers, it can be considered a measure of verbal fluency, not necessarily of creativity (Silvia et al., 2008). Overall, divergent thinking tasks measure only a very limited aspect of creativity, and therefore their relevance in education has been questioned (Barbot, Besancon, & Lubart, 2015; Zeng, Proctor, & Salvendy, 2011).

In addition to the measures of creativity as an individual's ability or potential (e.g. self-evaluations, divergent thinking and insight tasks), creativity can also be measured in a product, such as creativity of a novel. Assessing creativity of a product instead of focusing on individual's ability or potential may be particularly applicable in education (Barbot, Besancon, & Lubart, 2015). Creative outcome, such as a short story or picture, is not only a product of creative potential but also reflects other factors, such as intrinsic motivation and domain-specific skills (Amabile, 1983).

2. Problem Statement

Previous research on the relationship between creativity and educational achievement has focused on creative potential measures, namely divergent thinking (see meta-analysis, Gajda et al., 2016). Creative potential measures assess individual's ability, not creative behaviours or outcomes. This study will use Consensual Assessment Technique (CAT; Amabile, 1982) to estimate the creativity of a product, namely written stories. The use of CAT to evaluate creativity in written texts has demonstrated high inter-rater reliability in a previous study (Baer, Kaufman, & Gentile, 2004). Most of the previous studies on creativity and education have applied a cross-sectional design. The longitudinal nature of the current study is a particular strength. The availability of diverse measures in our study sample also allows us to investigate the role of creativity in educational achievement in addition to intelligence and personality. Also, the stories were written at children's homes, not in the school environment, which may influence the creative expressiveness in the stories.

3. Research Questions

1. Can CAT be used as a method to estimate creativity inchildren's written stories at age 9?
2. Does story creativityat age 9 predict educational achievement in English, and/or in Maths at age 16 over intelligence, previous school achievement and personality measures associated with creativity?

4. Purpose of the Study

This study will add incremental knowledge on the use of CAT to evaluate creativity in written children's stories. Furthermore, this study will investigate whether creativity in writing can already be detected in primary schooland whetherit relates to further educational achievement.

5. Research Methods

5.1. Sample

A randomly selected subsample of 60 twins from the Twins Early Development Study (TEDS) was used in this study. TEDS is a large, longitudinal twin sample that includes more than 13,000 twin pairs, born between 1994 and 1996, representative of the population of England and Wales (Haworth et al., 2013). Only one twin per pair was selected to eliminate the effect of the shared home environment. Data from one participant was excluded from further analyses due to poor handwriting that made it impossible to transcribe the story. The final sample consisted of 40 females and 19 males.

5.2. Measures

The measures were collected in two different data collection waves at ages 9 and 16.

5.2.1. Written stories at 9

Stories were written at age 9. The children were shown three coloured pictures of animals and buildings at a farm. They were then instructed to write a story that was creative. The pictures and instructions for the task are shown below in Figure 01. The data collection was done in children's homes. There was no time limit for the task and it was instructed and supervised by the parents. The stories were transcribed to minimise the influence of differences in handwriting in coding. No corrections to the stories were made in the transcription so the spelling mistakes were also included.
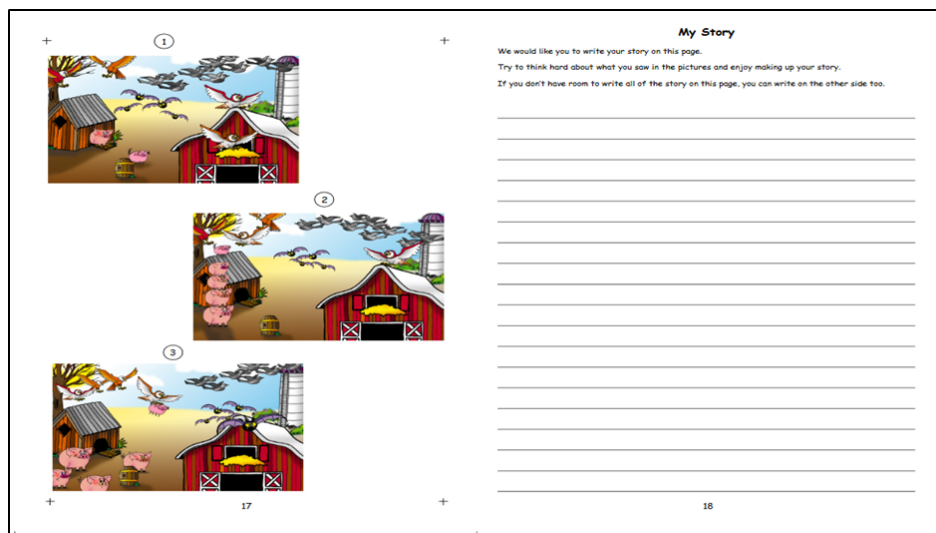
Figure 01. The pictures and instructions for the My Story task.

The Consensual Assessment Technique (Amabile, 1983) is widely used, reliable and validated creativity evaluation technique. It is based on the assumption that a group of independent judges are best able to make evaluations on the creativity of a product (Hennessey & Amabile, 1999). Creativity may be difficult to define and characterise, but as demonstrated by CAT, people can recognise and agree on it (Hennessey & Amabile, 1999). CAT has been used in different domains among primary school children,

such as on musical compositions and drawings (Hickey, 2001; Lubart et al., 2010). CAT has also previously been used to evaluate creativity in children's orally told stories (Hennessey & Amabile, 1988).

In the present study, six independent judges coded the stories for 10 different dimensions each on a 5-point Likert-scale using their own subjective interpretation of creativity. As instructed by the method of CAT, no fixed criteria were presented to judges on which their scoring should be based (Amabile, 1983). The judges were instructed to evaluate creativity in the stories as follows: "Please evaluate the creativity of the story on this page in relation to the other 58 stories. Use your own subjective assessment of creativity". Nine other dimensions that the judges were asked to evaluate in the stories were: Liking, Novelty, Imagination, Logic, Emotion, Grammar, Detail, Vocabulary and Straightforwardness (Hennessey & Amabile, 1988). Three of the judges were primary school teachers, three undergraduate Psychology students. All the judges were females.

5.2.2. General cognitive ability at 9

General cognitive ability at age 9 was a combination of two non-verbal tests and two verbal tests. The non-verbal Puzzle and Shapes tests are part of the Cognitive Abilities Test 3(CAT3; Smith, Fernandez, & Strand, 2001). Verbal ability at age 9 was assessed using the vocabulary and general knowledge tests (WISC-III-UK; Wechsler, 1992).

5.2.3. English and Maths at 9

English and Maths at age 9 were standardised teacher-reported scores, each based on three different evaluations per subject. The three components evaluated in English were Speaking and Listening, Reading, and Writing. In Maths they were Using and Applying Mathematics, Numbers, and Shapes, Space and Measures.

5.2.4. Openness to Experience and Extraversion at 16

Openness to Experience and Extraversion were measured as part of 30-item personality scale (Mullins-Sweatt et al., 2006) based on the Five-Factor Model of personality. Each personality factor was assessed by 6 items.

5.2.5. English and Maths at 16

English and Maths scores at 16 are based on the results of the General Certificate of Secondary Education (GCSE), a standardised end of the school exam in the UK. English grade is the average of Language and Literature; Maths grade is the average of Maths, Statistics and Additional Maths.

6. Findings

The inter-rater reliabilities between the six judges for all the story dimensions were high (α = .76 - .95). The total score for each of the ten dimensions was calculated as a sum of the scores from all six judges. All the total dimension scores were normally distributed. To establish clusters between the ten dimensions, a Principal Component Analysis (PCA) was run with a Varimax rotation. The Kaiser-Meyer-

Olkin measure was high (KMO = .90) and Bartlett's test of sphericity indicated that the correlations between the coded dimensions in the stories were sufficient for PCA ($\chi2(45) = 851.55$, $p < .001$). The factor loadings are presented in Table 01.

Table 01. The rotated factor loadings on a Principal Components Analysis (Varimax rotation) for 10 dimensions coded in the stories

Dimension	Factor 1	Factor 2
	Creative Expressiveness	**Logic**
Creativity	**.95**	.30
Imagination	**.95**	.22
Novelty	**.94**	.19
Liking	**.88**	.42
Detail	**.84**	.33
Emotion	**.80**	.35
Vocabulary	**.75**	.48
Straightforwardness	.13	**.95**
Logic	.39	**.87**
Grammar	.44	**.75**

Note.n = 59

The factor loadings revealed two separate factors: Creative Expressiveness and Logic. Factor scores for these two factors were created by combining the scores from the dimensions that had factor loading higher than 0.7 (see bolded valued in Table1 above). To investigate the relationships between all the study variables, bivariate correlations were run for the Creativity dimension (age 9); Creative Expressiveness and Logic factor (age 9); general cognitive ability (age 9); English and Maths (at age 9); Openness to Experience and Extraversion (age 16) and English and Maths GCSE grades (age 16). The correlations are presented in Table 02.

Table 02. Correlations between all the study measures (n=59).

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1.	**Creativity dimension score at 9**	1									
2.	**Creative Expressiveness factor score at 9**	.97**	1								
3.	**Logic factor score at 9**	.59**	.62**	1							
4.	**General cognitive ability at 9**	.18	.20	.21	1						
5.	**English at 9**	.34**	.36**	.39**	.41**	1					
6.	**Maths at 9**	.28*	.29*	.33**	..44**	.58**	1				
7.	**Extraversion at 16**	-.02	-.02	.03	.05	-.05	-.05	1			
8.	**Openness to Experience at 16**	-.07	-.08	-.12	-.01	-.08	-.11	.41**	1		
9.	**English at 16**	.41**	.45**	.37**	.42**	.49**	.41**	-.05	-.21	1	
10.	**Maths at 16**	.32*	.31*	.31*	.45**	.49**	.69**	-.09	-.29*	.63**	1

*** $p < .01$; * $p < .05$*

To investigate whether the Creative Expressiveness factor score explains variance in Maths and/or English at age 16, hierarchical linear regressions were run. General cognitive ability score at age 9 was entered as a first step into the regression model, followed by previous academic achievement at age 9 (English or Math) in the second step. In the third step, Openness to Experience and Extraversion scores at age 16 were entered to the model. In the fourth and last step, Creative Expressiveness and Logic factor scores were entered into the regression model predicting either English or Maths score at age 16. The results from the hierarchical regression predicting English achievement at age 16 are presented in Table 03, and for Maths achievement at age 16 in Table 04.

Table 03. Summary of Hierarchical Regression Analysis for Variables predicting English at age 16.

Variable	β	t	R	R^2	ΔR^2
Step 1			.42	.18	.18
G at 9	.42	3.54**			
Step2			.55	.30	.12
G at 9	.27	2.18*			
English at 9	.38	3.12**			
Step 3			.59	.33	.03
G at 9	.27	2.21*			
English at 9	.37	3.01**			
Openness at 16	-.20	-1.60			
Extraversion at 16	.04	0.32			
Step 4			.64	.40	.07
G at 9	.25	2.17*			
English at 9	.27	.27*			
Openness at 16	-.18	-1.51			
Extraversion at 16	.03	.28			
F1_Creativity at 9	.29	2.50*			
F2_Logic at 9	.01	.07			

Note. n = 59; *p< .05; **p< .01

The hierarchical multiple regression revealed that at the first step, general cognitive ability contributed significantly to the regression model ($F(1,58) = 12.52$, $p < .01$) and accounted for 18% of the variance in English results at age 16. Includingthe previous achievement measure, English at age 9, explained an additional 12% of the variance ($F(2,58) = 12.01$, $p < .01$). Personality measures of Openness to Experience and Extraversion were not individually significant predictors of English at 16. In the last step, adding Creative Expressiveness and Logic factor scores into the model explained an additional 7% of the variance was in the English score at 16 ($F(6,58) = 7.19$, $p < .01$). Only the Creative Expressiveness, and not the Logic, was a significant predictor.

Table 04. Summary of Hierarchical Regression Analysis for Variables predicting Maths at age 16.

Variable	β	t	R	R^2	ΔR^2
Step 1			.45	.21	.21
G at 9	.45	3.85**			
Step2			.71	.51	.30
G at 9	.18	1.76			
Maths at 9	.61	5.87**			
Step 3			.75	.56	.05
G at 9	.28	1.92			

Maths at 9	.94	5.75**			
Openness at 16	-.51	-2.34*			
Extraversion at 16	.05	.24			
Step 4			.75	.57	.01
G at 9	.19	1.81			
Maths at 9	.56	5.27**			
Openness at 16	-.23	-2.26*			
Extraversion at 16	.02	.23			
F1_Creativity at 9	.10	.86			
F2_Logic at 9	-.01	-.01			

Note. n = 59; *$p< .05$; **$p< .01$

In the regression model explaining variance in Maths grade at 16, general cognitive ability ($F_{(1,58)} = 12.52$, $p < .01$) accounted for 21% of the variation in Maths results at age 16. Including the Maths score at age 9, explained an additional 30% of the variation ($F_{(2,58)} = 12.01$, $p < .01$). Out of the two personality measures, only Openness to Experience was a significant (negative) predictor of Maths score at 16 ($F_{(1,58)} = 12.52$, $p < .01$). Creative Expressiveness and Logic factor scores did not explain any additional variance in the model when added in the last step.

7. Conclusion

The present study set to test the robustness of the Consensual Assessment Technique as a method to evaluate children's written stories at age 9. Our results showed that CAT is a reliable measure of creativity in the children's written stories. The inter-rater reliability for the Creativity dimension was the highest out of all coded ten dimensions. Further factor analysis on the ten dimensions revealed that Creativity loaded onto a single factor along with Imagination, Novelty, Liking, Detail, Emotion and Vocabulary (Creative Expressiveness factor). Straightforwardness, Grammar and Logic formed a second, separate factor (Logic factor).

Additionally, we aimed to investigate whether creativity, measured by CAT, is associated with later educational achievement at age 16. Our results showed that the Creative Expressiveness factor score explained an additional 7% of variance in English GCSE grade at age 16, above and beyond intelligence and English grade measured at age 9. This suggests that marking criteria in English at age 16 includes both, technical knowledge(spelling, grammar, etc.), as reflected in the English grade at age 9; and creativity (explicitly or implicitly), captured by our Creative Expressiveness measure.Our results suggest that creativity in writing is overlooked in the UK primary education marking criteria, as creativity was not captured by the English grade at age 9. It is possible that primary education focuses more on convergent and technical skills, or at least did at the time of the data collection (the stories were written in 2003-2005).

Future studies are needed to explore how to apply the CAT in the evaluation of creative expressiveness in writing, as well as other domains, such as music and arts, in children of different ages. Further research is needed to establish how the evaluation of creativity in primary education, as guided by CAT, can be done in practice. Having independent judges evaluating children's work may not be easy to organise, but the observed high level of agreement among judges suggests that, with some calibration,

single judges may provide accurate evaluations. Additionally, using a bigger sample, we plan to investigatewhether intrinsic motivation and/or enjoyment of writing moderate the relationship between creative expressiveness in written stories and further educational achievement. In conclusion, CAT is a promising direction for measuring creativity in children, with implications for both creativity research and educational practice.

Acknowledgments

References

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology*, *45*(2), 357.

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*(4), 309-313.

Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the consensual assessment technique to nonparallel creative products. *Creativity research journal*, *16*(1), 113-117.

Barbot, B., Besançon, M., & Lubart, T. (2015). Creative potential in educational settings: Its nature, measure, and nurture. *Education 3-13*, *43*(4), 371-381.

Beaty, R. E., Nusbaum, E. C., & Silvia, P. J. (2014). Does insight problem solving predict real-world creativity?. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(3), 287.

Besançon, M., & Lubart, T. (2008). Differences in the development of creative competencies in children schooled in diverse learning environments. *Learning and individual differences*, *18*(4), 381-389.

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in cognitive sciences*, *9*(7), 322-328.

Chamorro-Premuzic, T., & Reichenbacher, L. (2008). Effects of personality and threat of evaluation on divergent and convergent thinking. *Journal of Research in Personality*, *42*(4), 1095-1101.

Craft, A. (2003). The limits to creativity in education: Dilemmas for the educator. *British journal of educational studies*, *51*(2), 113-127.

Cropley, A. (2006). In praise of convergent thinking. *Creativity research journal*, *18*(3), 391-404.

Cropley, A. J. (1967). Creativity: a new kind of intellect? *Australian Journal of Education*, *11*(2), 120-125.

Dirkes, M. A. (1978). The role of divergent production in the learning process. *American Psychologist, 33*(9), 815-820.

Gajda, A., Karwowski, M., & Beghetto, R. A. (2016). Creativity and Academic Achievement: A Meta-Analysis.*Journal of Educational Psychology, 109*(2), 260-299.

Guilford, J. P. (1957). Creative abilities in the arts. *Psychological Review, 64*(2), 110-118.

Haworth, C. M., Davis, O. S., & Plomin, R. (2013). Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood. *Twin Research and Human Genetics*, *16*(1), 117-125.

Hennessey, B. A., & Amabile, T. M. (1988). Story-telling: A method for assessing children's creativity. *The Journal of Creative Behavior*, *22*(4), 235-246.

Hennessey, B. A., & Amabile, T. M. (1999). Consensual assessment. *Encyclopedia of creativity*, *1*, 347-359.

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*(3), 234-244.

Jeffrey, B. (2006). Creative teaching and learning: towards a common discourse and practice. *Cambridge Journal of Education*, *36*(3), 399-414.

Lubart, T., Pacteau, C., Jacquet, A. Y., & Caroff, X. (2010). Children's creative potential: An empirical study of measurement issues. *Learning and Individual Differences*, *20*(4), 388-392.

Mullins-Sweatt, S. N., Jamerson, J. E., Samuel, D. B., Olson, D. R., Widiger, T. A. (2006). Psychometric Properties of an Abbreviated Instrument of the Five-Factor Model. Assessment, 13: 119-137.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational psychologist*, *39*(2), 83-96.

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, *24*(1), 66-75.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., . . . Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts, 2*(2), 68-85.

Smith, P., Fernandez, C., & Strand, S. (2001). Cognitive abilities test 3 (CAT3). Windsor:nferNELSON.

Torrance, E. P., Ball, O. E., & Safter, H. T. (2003). *Torrance tests of creative thinking*. Scholastic Testing Service.

Wilson, A. (Ed.). (2014). *Creativity in primary education*. Learning Matters.

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*(1), 24-37.