Personality and Individual Differences 106 (2017) 36–40

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

From Likert scales to images: Validating a novel creativity measure with image based response scales

Franziska Leutner [a,*], Adam Yearsley [b], Sonia-Cristina Codreanu [a], Yossi Borenstein [c], Gorkan Ahmetoglu [a]

[a] Department of Psychology, UCL, 26 Bedford Way, London WC1H 0AP, UK
[b] Red Bull, Am Brunnen 1, 5330 Fuschl am See, Austria
[c] Independent Data Science Consultant, 8 Mansfield Road, NW32HN London, UK

ARTICLE INFO

Article history:
Received 31 May 2016
Received in revised form 6 October 2016
Accepted 8 October 2016
Available online 22 October 2016

Keywords:
Creativity
Image-based measure
Psychological assessment
Openness

ABSTRACT

The use of image-based testing to assess individual differences has increased substantially in recent years, with proponents arguing that they offer a more engaging alternative to text-based psychometric tests. Yet research examining the validity of these tests is near to non-existent. Traditional image-based formats have been little more than an adaptation of self-reports, with images replacing questions but not response options. The current study develops a novel image-based creativity measure, where images replace conventional response scales, and scores on the measures are obtained using a linear regression scoring algorithm to predict three self-reported creativity measures. Using sequential forward selection on a set of 77 image-based items, an optimal solution of 14 items that were valid predictors of self-reported creativity scores were identified. The image-based measure had good test-retest reliability. Implications are discussed in terms of the usefulness of image-based testing for practitioners seeking engaging and short test formats.

1. Introduction

The assessment of individual differences in psychological traits, such as personality, intelligence, and creativity, stretches back more than a century (Chamorro-Premuzic, 2007). The most common way of measuring differences between people is through psychometric tests (Ahmetoglu & Chamorro-Premuzic, 2013). Psychometric tests are used extensively in settings from selection (Rothstein & Goffin, 2006) to psychiatric diagnosis (Gilbody, Richards, Brealey, & Hewitt, 2007), and consumer profiling (Matz, Gladstone, & Stillwell, 2016). Despite their widespread use, psychometric tests are criticised for their inability to engage the test taker (Krosnick, 1991), the ease of faking responses (Morgenson et al., 2007), and adverse impact (Hough, Oswald, & Ployhart, 2001).

Perhaps in response to these criticisms, and fuelled by technological advances, recent years have seen mounting interest in more engaging forms of assessment (Attali & Arieli-Attali, 2015), including gamification (Chamorro-Premuzic & Steinmetz, 2013; Landers & Callan, 2011; Reeves & Read, 2013) and social media analytics (Kosinski, Matz, & Gosling, 2015; Pennebaker, 2011). However, innovative assessment tools often serve entertainment purposes, with little indication to their validity (Naglieri et al., 2004). The increase in the quantity of these instruments has not been synonymous with an increase in research into their quality, that is, their reliability and validity. Indeed, the desire to use innovative assessment by professionals has outpaced the peer-reviewed literature (e.g., Roth, Bobko, Van Iddekinge, & Thatcher, 2013). This gap between research and practise is problematic if tests are used to make hiring decisions or provide clinical diagnosis.

Consequently, developing scientific evidence for the validity and utility of image-based tests is critical, not only from an academic, but also an applied perspective. The current study takes a step in this direction. Specifically, an image-based creativity assessment and a predictive scoring algorithm are developed. The test-retest reliability, as well as its concurrent validity in relation to three text-based, self-report creativity measures are assessed, so that practitioners may better understand how such image-based tests compare to traditional tests.

1.1. Advantages of image-based formats

One of the most common innovations in psychological assessment formats has been to replace item questions with visual representations, thereby increasing user engagement (Barrett & Ebbeling, 2003; Downes-Le Guin, Baker, Mechling, & Ruylea, 2012; Hamari, Koivisto, & Sarsa, 2014; Lugtigheid & Rathod, 2005). Beyond engagement, image-based formats could provide theoretical and practical advantages over text-based psychometric tests. First, they may be more suitable for culturally and linguistically diverse test takers, and remove misunderstanding of text items (Paunonen, Jackson, & Keinonen, 1990). Second, responding to image-based items may require less attention, reducing test taker fatigue. Finally, image stimuli evoke stronger preferences in respondents than verbal stimuli, providing for reduced length of image-based tests (Lugtigheid & Rathod, 2005; Meissner & Rothermund, 2015).

* Corresponding author.
E-mail address: f.leutner.12@ucl.ac.uk (F. Leutner).

http://dx.doi.org/10.1016/j.paid.2016.10.007

1.2. Past research on image-based tests

Despite being innovative, image-based formats in assessment are not new. Geist's (1959) Pictorial Interest Inventory pictures a person engaged in three activities, of which respondents pick the most appealing one. More recent image-based tests adapt text-based personality measures such that the question is replaced with an image: The Nonverbal Personality Questionnaire (Paunonen et al., 1990) measures Murray's (1938) psychological needs such that participants report the likelihood that they would engage in visually displayed behaviours. A version of the test measuring the Big Five also exists (Paunonen, Ashton, & Jackson, 2001).

These adaptations of verbal personality tests have gained support in the academic literature for their internal reliabilities and validities (Hong, Paunonen, & Slade, 2008; Moore, Schermer, Paunonen, & Vernon, 2010; Paunonen, 2003; Paunonen, Jackson, Trzebinski, & Forsterling, 1992; Paunonen, Zeidner, Engvik, Oosterveld, & Maliphant, 2000). However, research examining the validity of image-based tests is scarce, and their use is mostly limited to special populations, such as children or illiterates. In addition, the use of response scales and scoring methodologies developed for verbal formats is not ideal: by using images to replace the question stem, questions are limited to those that can be visually represented.

1.3. Assessment of creativity

Creativity encompasses both personality and cognitive aspects related to the production of unique and useful ideas (Runco & Jaeger, 2012; Simonton, 2000). Three of the many components associated with creativity are: Cognitive Flexibility, the ability to switch cognitive sets to adapt to changing environmental stimuli (Scott, 1962); Curiosity, the recognition, pursuit, and intense desire to explore novel and uncertain events (Kashdan & Silvia, 2009) and Openness to Experience, the Big Five personality trait considered as a proxy of creativity (Feist, 1998; Furnham & Bachtiar, 2008; Martindale, 1989).

Because of the broadness of the construct, multi-trait, multi-method approaches have been proposed as most suitable (Cropley, 2000; Plucker & Makel, 2010). An image-based measure of creativity may add to this array of measurement methodologies available for creativity testing. In addition, image-based response scales may be particularly effective in measuring creativity because images elicit aesthetic preferences, such as preferences for complexity, which in turn are indicative of self-reported creativity and aesthetic styles (Barron, 1953; Chamorro-Premuzic, Reimers, Hsu, & Ahmetoglu, 2009; Rawlings, 2003; Swami, Stieger, Pietschnig, & Voracek, 2010; Wiersema, van der Schalk, & van Kleef, 2012). A preference for complex polygons is associated with higher self-reported creativity, such that Eisenman and Robinson (1967, 1968) suggested the use of polygons varying in their level of complexity as measures of creativity. Accordingly, the present research aimed to a) develop a novel format image-based creativity measure, b) investigate its concurrent validity in relation to three text-based measures of creativity, and c) assess its test-retest reliability.

2. Method

2.1. Measures

2.1.1. Curiosity and Exploration Inventory-II (CEI-II; Kashdan et al., 2009)

A 10-item, five-point Likert self-report scale. The CEI-II measures two traits: stretching (e.g., 'I actively seek as much information as I can in new situations') and embracing (e.g., 'I am the type of person who really enjoys the uncertainty of everyday life'). The CEI-II demonstrates reliability estimates of 0.85, construct validity, discrimination, desirable breadth of difficulty (Kashdan et al., 2009), and predictive validity for task performance (Kashdan, Rose, & Fincham, 2004).

2.1.2. Cognitive Flexibility Inventory (CFI; Dennis & Vander Wal, 2010)

A 20-item, seven-point Likert scale, self-report measure of adaptive thinking in stressful situations. Thirteen items assess behaviours related to alternatives (e.g., 'I consider multiple options before making a decision'), and seven items behaviours related to control ('When I encounter difficult situations, I feel like I am losing control'). The CFI shows a reliable factor structure, internal consistency, test-retest reliability, and concurrent validity (Dennis & Vander Wal, 2010).

2.1.3. Openness to experience (Goldberg, 1999)

Measured on a five-point Likert scale ('very inaccurate' to 'very accurate') using the 10-item Openness scale from the International Personality Item Pool (e.g. 'I enjoy hearing new ideas').

2.2. Item design

The question stem of image-based items retained its verbal format, but the response scale presented a range of images (see Fig. 1). Each item consisted of a text-based question and between two and eight image response options. The image response options took one of two forms: they either assessed varying levels of the same trait, or they represented different traits. Seventy-seven items were designed to reflect Cognitive Flexibility, Curiosity, and Openness.

2.3. Scoring

The scoring algorithm was developed on a sample of 964 participants, recruited using a UK panel company, and compensated for their participation. The panel had an equal distribution of males and females, and participants were UK residents. Approximately half of the users were 18–25 and the other half 25–36 years old. Participants completed the three creativity measures as well as all 77 image-based items.

Rather than stipulating which responses were indicative of which underlying trait, responses to image-based items were scored in relation to standard measures. This method is commonly used in measure validation procedures when testing concurrent validity between new and existing measures (Rust & Golombok, 2009), as well as for predictive personality measures (Bachrach, Kosinski, Graepel, Kohli, & Stillwell, 2012; Boyd et al., 2015; Lambiotte & Kosinski, 2014; Youyou,

Fig. 1. Example image response scales.

Kosinski, & Stillwell, 2015; Schwartz et al., 2013; Wang, Kosinski, Stillwell, & Rust, 2014).

Responses to all 77 items were dummified, with each image response option being transformed into a binary variable. This resulted in 321 dummy variables. Dummified responses were used as the independent variables (predictors) and the creativity scores as the respective dependent (predicted) variables in linear regression models to estimate the creativity scores.

2.3.1. Item selection

As the large number of dimensions resulting from 321 dummy variables can cause over-fitting, two methods of feature selection were used. The first method, LASSO (Least Absolute Shrinkage and Selection Operator) regression with 10-fold cross validation was applied to reduce the number of image response options. LASSO is a regularized regression, which penalizes variables with large coefficients and discounts variables with inconsistent performance across the sample. Thereby LASSO selects response options that are most indicative of creativity.

LASSO cannot take into account that some image response options were taken from the same question. In order to account for the contribution of single questions, a second feature selection method, Sequential Forward Selection, was used (Devijver & Kittler, 1982). Starting with an empty set of questions, LASSO regression with 10-fold cross validation was used to estimate the relevant scale. The predicted and measured scores were correlated, and additional questions added at each step until no new question improved the correlation by >0.1. Questions with individual correlations higher than 0.2 were also retained. This resulted in a final set of 14 questions, or 64 dummy variables.

2.4. Results

With the selected 14 questions as predictor variables, LASSO regression with 10-fold cross validation was performed to predict creativity scores. Coefficients for the models predicting Cognitive Flexibility, Curiosity, and Openness, are presented in Table 1.

2.5. Validation

1071 participants (605 females) were recruited using Amazon's Mechanical Turk (MTurk). Participants completed the text-based creativity measures and the 14-item image-based measure, which is part of the Red Bull Wingfinder assessment. To assess test-retest reliability, a subset of 162 participants retook the test after 60 days. MTurk panellists were US citizens paid for their participation. 15% were aged 18–24, 47% aged 25–34, 24% aged 35–44, and 14% aged 45 to 59. Responses to the image-based measure were scored using the algorithm described in Section 2.3.

Creativity scores were normally distributed on the text- and image-based measures (see Table 2). The three text-based creativity scores had moderate intercorrelations (average $r = 0.47$, with $p < 0.001$), as had the three image-based scores (average $r = 0.5$, with $p < 0.001$) (see Table 2).

Correlations between text- and image-based scores were moderate to high. Concurrent validity was higher for Curiosity and Openness than for Cognitive Flexibility (see Fig. 2). The average test-retest reliability of the image-based measures was $r = 0.63$ ($p < 0.001$) (see Fig. 2).

3. Discussion

The aim of this study was to examine the psychometric properties of a newly developed image-based creativity measure. Results obtained from two large samples provided preliminary evidence for the test-retest reliability and concurrent validity of the 14-item measure.

The developed scoring algorithm accurately predicted creativity scores on two of the three existing scales. This finding is in line with studies demonstrating the use of predictive models for measuring personality (Chen, Hsieh, Mahmud, & Nichols, 2014; Lambiotte & Kosinski, 2014; Yarkoni, 2010) and indicates that predictive scoring algorithms are suitable for scoring image-based response scales.

Table 1
Model coefficients for predicting creativity from image response options.

	Cognitive Flexibility	Curiosity	Openness
Intercept	0.43	0.59	0.44
How do you feel when plans change?		0.0420	
		0.0000	
How do you like to think?			−0.0366
			0.0070
How many good solutions exist to a problem?	0.0000		
	0.0214		
	0.0009		
	−0.0416		
	0.0000		
What do you spend the most time thinking about?	0.0478	0.0289	
	0.0000	0.0000	
	−0.0298	−0.0463	
Which is more like you?		0.0406	
		−0.0038	
Which is most like you?			0.0585
			0.0142
			−0.0009
			−0.0033
Which kinds of books do you never read? Pick as many as you need-			−0.0339
			−0.0114
			−0.0239
			−0.0185
			−0.0592
			−0.0102
			0.0012
			−0.0071
Which of these do you feel strongly about? Pick as many as you like-			0.0626
			0.0221
			0.0106
			0.0158
Which of these have you been to in the last three weeks? Pick as many as you need-		0.0199	0.0601
		0.0594	0.0020
		0.0164	0.0082
		−0.0635	−0.0120
		0.0240	−0.0080
		0.0140	−0.0203
		−0.0101	−0.0462
		−0.0064	0.0170
Which of these have you done in the last three weeks? Pick as many as you need-	0.0358		0.0273
	0.0000		0.0010
	0.0002		0.0193
	−0.0023		0.0281
	0.0000		0.0010
	0.0056		−0.0089
	0.0039		0.0178
	0.0143		0.0217
Which of these represents an opportunity for learning? Pick as many as you like-	0.0006		0.0020
	0.0000		0.0210
	0.0000		−0.0047
	0.0101		0.0019
	0.0000		−0.0104
	0.0000		0.0197
	0.0131		0.0316
	0.0399		0.0088
You're visiting a new country—How immersed do you get in local culture?	0.0453	0.0610	0.0404
	0.0000	0.0000	0.0000
	−0.0014	−0.0931	−0.0671
	−0.0090	−0.0692	−0.0398
You're visiting a new country—How immersed do you get in local culture?		0.0003	
		0.0310	
		−0.0279	
		−0.0584	
You've been offered a free subscription to a magazine of your choice—Which topic would you choose?	−0.0015		
	0.0000		
	−0.0003		
	0.0000		
	0.0000		
	0.0000		

Table 2
Descriptive statistics and correlations for image- and text-based measures.

		Pearson correlation coefficient							Mean	SD
		Text-based measure			Image-based measure			Gender		
		Cog Flex	Curiosity	Openness	Cog Flex	Curiosity	Openness			
Text-based measure	Cog Flex								97.3	16
	Curiosity	0.48							31.65	7.89
	Openness	0.41	0.44						38.96	6.52
Image-based measure	Cog Flex	0.35	0.41	0.28					0.59	0.09
	Curiosity	0.32	0.5	0.27	0.54				0.57	0.11
	Openness	0.25	0.33	0.51	0.49	0.48			97.3	16
	Gender	−0.04	−0.08	0.07	0.12	−0.05	0.16		1.55	0.51
	Age	0.1	−0.12	−0.02	−0.01	−0.01	0.05	0.08	4.57	2.19

Notes: Cog Flex = Cognitive Flexibility. Female = 2, male = 1. Age groups: 1 = 18–20, 2 = 21–24, 3 = 25–29, 5 = 30–34, 6 = 35–39, 7 = 40–44, 8 = 45–49, 9 = 50–54, 10 = 55–59, 11 = 60–64, 12 = 65+. Correlations > 0.1 (>0.05) are significant at the $p < 0.01$ (0.05) level.

Moderate correlations between the image-based and the text-based measures for Curiosity and Openness demonstrated good concurrent validity of the image-based format ($r = 0.5$, $p < 0.001$). Furthermore, the measure exhibited good test retest reliability ($r = 0.65$, $p < 0.001$), indicating that the selected image-based items are able to reliably measure aspects of creativity. On the other hand, the concurrent validity for Cognitive Flexibility was relatively low ($r = 0.35$, $p < 0.001$), suggesting that the selected images may not assess this particular aspect of creativity equally well.

The predictive scoring algorithm used fewer items than established text-based measures, assessing all three creativity aspects with 14 items, compared to 40 items on the text-based measures. Both the predictive scoring algorithm and stronger associations evoked by images may be reasons for achieving shorter length (Meissner & Rothermund, 2015).

The image-based measure demonstrated good test-retest reliability (average $r = 0.63$, $p < 0.001$), in particular taking into account factors that might have reduced the correlation including the small number of items, long interval between test and retest (six weeks), and the small to moderate sample size. Indeed, the observed test-retest reliability for the image-based Openness measure was higher than that reported in other studies for the ten-item, text-based Openness measure (reported $r = 0.55$ in Kosinski, Stillwell, & Graepel, 2013). Participants were more likely to consistently select the same image than they were to consistently select the same point on a Likert scale. This could be due to the relatively broader construct of creativity as compared to Openness (i.e. broader constructs tend to display better reliability; Chamorro-Premuzic, 2011). In addition, some image-based items had only two images as response options, compared with five to seven response options on Likert scales, which could lower the probability of changing responses.

3.1. Implications

The current study has a number of implications for the development of image-based assessments, particularly those focusing on creativity, but also beyond. For researchers and practitioners interested in alternatives to traditional self-report Likert response scales, this study provides preliminary support for the validity and reliability of an image-based measure. Although additional research is needed to replicate and extend these findings, this study takes a step towards providing evidence for the utility of innovative psychological assessments.

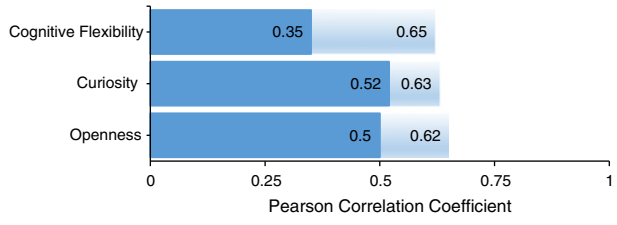

Fig. 2. Concurrent validity of the image-based measure expressed by correlations between image-based score and text-based score. Transparent bars indicate the baseline accuracy of the image-based measure expressed by test-retest correlation of image-based scores. Correlations are significant at $p < 0.001$.

The image-based measure has a number of advantages in practice. The measure is shorter than both existing text- and image-based creativity measures. Image response scales may be less obvious in what they are measuring than Likert scales. As a consequence, image-based scales could be less prone to faking and appear less intrusive to the test taker.

3.2. Limitations and future research

The current study has a number of limitations. It assessed the concurrent validity of image-based measures in relation to text-based, self-report inventories only. Although this provides initial evidence of the validity of image-based measures, additional studies are needed to investigate its relationship to non-self-report measures of creativity, such as divergent thinking tests. In addition, reliability of the image-based measure should be improved by identifying additional image-based items, in particular if the measure is to be used in selection contexts. The incremental validity of the image-based measures in predicting creative performance, beyond existing tests, should be established. This research would be necessary for demonstrating the value of using image-based measures alongside, in addition to, or as a replacement of, current creativity measures.

There has been a call for a multi-method, multi-trait approach to the study of creativity (Batey & Furnham, 2006; Cropley, 2000; Park, Chun, & Lee, 2016; Plucker & Makel, 2010). Image-based measures may tap into the same performance variance as other creativity measures. But provided that image-based measures are more engaging than self-reports and easier to administer than divergent thinking tests, they could provide an alternative to existing creativity measures. Otherwise, image-based tests may predict distinct performance variance from current measures. In this scenario, they could be used alongside traditional tests. Accordingly, the purpose of developing an image based assessment is not only to provide more engaging alternatives for established (self-report) methods, but to provide an optional methodology for a valid multi-method approach to assessing creativity (and perhaps other personality traits like the Big Five).

3.3. Conclusion

This study supports the proposition that creativity can be measured via preferences for image-based stimuli. It may encourage research into innovative assessment formats and help practitioners in applying alternative assessments in settings where evidence of validity and reliability is required. Image-based assessments may provide a solution to an

evolving need for alternative assessments, and this study was one of the first to attempt to bridge the gap between practise and research.

References

Ahmetoglu, G., & Chamorro-Premuzic, T. (2013). *Personality 101*. New York: Springer Publishing Company.

Attali, Y., & Arieli-Attali, M. (2015). Gamification in assessment: Do points affect test performance? *Computers & Education*, *83*, 57–63.

Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012, June). Personality and patterns of Facebook usage. *Proceedings of the 4th Annual ACM Web Science Conference* (pp. 24–32). ACM.

Barrett, P., & Ebbeling, N. (2003). *Personality assessment via a graphical profiler. Paper presented at NZ Annual Psychology Conference*. New Zealand: Palmerston North.

Barron, F. (1953). Complexity-simplicity as a personality dimension. *The Journal of Abnormal and Social Psychology*, *48*(2), 163–172.

Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, *132*(4), 355–429.

Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., & Mihalcea, R. (2015). Values in words: Using language to evaluate and understand personal values. *Proceedings of the Ninth International AAAI Conference on Web and Social Media* (pp. 31–40).

Chamorro-Premuzic, T. (2007). *Personality and individual differences*. Malden: Blackwell Publishing.

Chamorro-Premuzic, T. (2011). *Personality and individual differences*. Chichester: Blackwell Publishing.

Chamorro-Premuzic, T., & Steinmetz, C. (2013). The perfect hire. *Scientific American Mind*, *24*(3), 42–47.

Chamorro-Premuzic, T., Reimers, S., Hsu, A., & Ahmetoglu, G. (2009). Who art thou? Personality predictors of artistic preferences in a large UK sample: The importance of openness. *British Journal of Psychology*, *100*(3), 501–516.

Chen, J., Hsieh, G., Mahmud, J., & Nichols, J. (2014). Understanding individuals' personal values from social media word use. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing. February*. (pp. 405–414).

Cropley, A. (2000). Defining and measuring creativity: Are creativity tests worth using? *Roeper Review*, *23*, 72–79.

Dennis, J. P., & Vander Wal, J. S. (2010). The cognitive flexibility inventory: Instrument development and estimates of reliability and validity. *Cognitive Therapy and Research*, *34*(3), 241–253.

Devijver, P. A., & Kittler, J. (1982). *Pattern recognition: A statistical approach*. Englewood Cliffs: Prentice Hall.

Downes-Le Guin, T., Baker, R., Mechling, J., & Ruylea, E. (2012). Myths and realities of respondent engagement in online surveys. *International Journal of Market Research*, *54*(5), 1–21.

Eisenman, R., & Robinson, N. (1967). Complexity-simplicity, creativity, intelligence, and other correlates. *The Journal of Psychology*, *67*(2), 331–334.

Eisenman, R., & Robinson, N. (1968). Peer-, self-, and test-ratings of creativity. *Psychological Reports*, *23*(2), 471–474.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, *2*(4), 290–309.

Furnham, A., & Bachtiar, V. (2008). Personality and intelligence as predictors of creativity. *Personality and Individual Differences*, *45*(7), 613–617.

Geist, H. (1959). The Geist picture interest inventory: General form: Male. *Psychological Reports*, *5*, 413–438.

Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): A diagnostic meta-analysis. *Journal of General Internal Medicine*, *22*(11), 1596–1602.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe. 7(1)*. (pp. 7–28).

Hamari, J., Koivisto, J., & Sarsa, H. (2014, January). Does gamification work?—A literature review of empirical studies on gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences*, Hawaii, USA, January 6–9, 2014.

Hong, R. Y., Paunonen, S. V., & Slade, H. P. (2008). Big Five personality factors and the prediction of behavior: A multitrait–multimethod approach. *Personality and Individual Differences*, *45*(2), 160–166.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, *9*(1–2), 152–194.

Kashdan, T. B., & Silvia, P. J. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. *Oxford handbook of positive psychology. 2*. (pp. 367–374).

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment*, *82*(3), 291–305.

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, *43*(6), 987–998.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Science of the United States of America. 110(15)*. (pp. 5802–5805).

Kosinski, M., Matz, S. C., & Gosling, S. D. (2015). Facebook as a research tool for the social sciences. *American Psychologist*, *70*(6), 543–556.

Krosnick, J. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*, 213–236.

Lambiotte, R., & Kosinski, M. (2014). Tracking the digital footprints of personality. *Proceedings of the IEEE*, *102*(12), 1934–1939.

Landers, R. N., & Callan, R. C. (2011). Casual social games as serious games: The psychology of gamification in undergraduate education and employee training. *Serious games and edutainment applications* (pp. 399–423). London: Springer.

Lugtigheid, A., & Rathod, S. (2005). *Questionnaire length and response quality: Myth or reality?* Stamford, CT: Survey Sampling International.

Martindale, C. (1989). Personality, situation, and creativity. In J. A. Glover, R. R. Ronning, & C. Reynolds (Eds.), *Handbook of creativity* (pp. 211–233). New York, London: Plenum Press.

Matz, S. C., Gladstone, J. J., & Stillwell, D. (2016). Money buys happiness when spending fits our personality. *Psychological Science*, *1*, 1–11.

Meissner, F., & Rothermund, K. (2015). A thousand words are worth more than a picture? The effects of stimulus modality on the implicit association test. *Social Psychological and Personality Science*, *6*(7), 740–748.

Moore, M., Schermer, J. A., Paunonen, S. V., & Vernon, P. A. (2010). Genetic and environmental influences on verbal and nonverbal measures of the Big Five. *Personality and Individual Differences*, *48*(8), 884–888.

Morgenson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, *60*(3), 683–729.

Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.

Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., & Velasquez, R. (2004). Psychological testing on the internet: New problems, old issues. *American Psychologist*, *59*(3), 150–162.

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), 1–10.

Paunonen, S. V. (2003). Big Five factors of personality and replicated predictions of behavior. *Journal of Personality and Social Psychology*, *84*(2), 411.

Paunonen, S. V., Jackson, D. N., & Keinonen, M. (1990). The structured nonverbal assessment of personality. *Journal of Personality*, *58*, 481–502.

Paunonen, S. V., Jackson, D. N., Trzebinski, J., & Forsterling, F. (1992). Personality structure across cultures: A multimethod evaluation. *Journal of Personality and Social Psychology*, *62*(3), 447.

Paunonen, S. V., Zeidner, M., Engvik, H. A., Oosterveld, P., & Maliphant, R. (2000). The nonverbal assessment of personality in five cultures. *Journal of Cross-Cultural Psychology*, *31*(2), 220–239.

Paunonen, S. V., Ashton, M. C., & Jackson, D. N. (2001). Nonverbal assessment of the Big Five personality factors. *European Journal of Personality*, *15*(1), 3–18.

Pennebaker, J. W. (2011). Your use of pronouns reveals your personality. *Harvard Business Review*, *89*, 32–33.

Plucker, J. A., & Makel, M. C. (2010). Assessment of creativity. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity*. New York, NY: Cambridge University Press.

Rawlings, D. (2003). Personality correlates of liking for 'unpleasant' paintings and photographs. *Personality and Individual Differences*, *34*(3), 395–410.

Reeves, B., & Read, J. L. (2013). *Total engagement: How games and virtual worlds are changing the way people work and businesses compete*. Harvard Business Press.

Roth, P. L., Bobko, P., Van Iddekinge, C. H., & Thatcher, J. B. (2013). Social media in employee-selection-related decisions a research agenda for uncharted territory. *Journal of Management*, *42*(1), 269–298.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review*, *16*(2), 155–180.

Runco, M. a., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*(1), 92–96.

Rust, J., & Golombok, S. (2009). *Modern psychometrics* (3rd ed). Hove: Routledge.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS One*, *8*(9), e73791.

Scott, W. A. (1962). Cognitive complexity and cognitive flexibility. *Sociometry*, 405–414.

Simonton, D. K. (2000). Creativity: Cognitive, personal, developmental, and social aspects. *American Psychologist*, *55*(4), 407–421.

Swami, V., Stieger, S., Pietschnig, J., & Voracek, M. (2010). The disinterested play of thought: Individual differences and preference for surrealist motion pictures. *Personality and Individual Differences*, *48*(7), 855–859.

Wang, N., Kosinski, M., Stillwell, D. J., & Rust, J. (2014). Can well-being be measured using Facebook status updates? Validation of Facebook's gross national happiness index. *Social Indicators Research*, *115*(1), 483–491.

Wiersema, D. V., Van Der Schalk, J., & van Kleef, G. A. (2012). Who's afraid of red, yellow, and blue? Need for cognitive closure predicts aesthetic preferences. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(2), 168–174.

Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, *44*(3), 363–373.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, *112*(4), 1036–1040.