Acta Psychologica 231 (2022) 103797

Contents lists available at ScienceDirect

Acta Psychologica

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

Decreases in divergent thinking across age groups from 2005 to 2018 amongst school children in Sudan

Salaheldin Fararh Attallah Bakhiet [a], Edward Dutton [b], Hatim Abdelrahman Warrag Ali [c], Guy Madison [e,*], Homoud Abdullah Saad Almoghyrah [a], Sultan Howedey Sultan Al-Mutairy [d], Yossry Ahmed Sayed Essa [a], Abdulelah Abdullah Saleh alruwaitea [a], Abdulrahman Saad Rashd Aljbr [a]

[a] King Saud University, Department of special Education, College of Education, Riyadh, Saudi Arabia
[b] Asbiro University, Lodz, Poland
[c] University of Khartoum, Khartoum, Sudan
[d] King Saud University, College of Education, Department of Educational Technology, Riyadh, Saudi Arabia
[e] Umeå University, Umeå, Sweden

ARTICLE INFO

Keywords:
Creativity
Flynn effect
Negative Flynn effect
Torrance
Sudan

ABSTRACT

Performances on tests of creativity have been found to be in decline in the USA. Here, we explore scores on divergent thinking tests in private schools in Khartoum State in Sudan by comparing a 2005 and a 2018 administration of the Torrance Standardized Circles test to 8- to 12-year-olds of both sexes. We find a decline across the period in all three dimensions of the test (Fluency, Flexibility and Originality), as well as in the overall index of divergent thinking. In line with much previous research, females consistently outperform males. Examining previous studies that report Negative Flynn Effects on IQ in Arab countries, we conclude that our results most likely reflect highly localized and exclusively environmental causes, and caution against assuming that the same processes that underlie Negative Flynn Effects in the West, whether on IQ or any trait correlated with it, also underlie it in the Arab World.

1. Introduction

Creativity is commonly defined as the ability to generate original ideas via the imagination in order to create or invent something (e.g. Sawyer, 2006, p. 287). The concept of creativity has provoked considerable discussion amongst students of individual differences, not least concerning how it is validly and reliably measured. There are two main approaches. One is to measure creative achievement, that is, to actually create something, with the provision that what is created also has some value or utility according to others in that domain (Jauk, 2019). The other is to assess divergent thinking: the ability to have lots of original and unusual ideas.

Both approaches are problematic. Measuring creative achievement requires that the person has actually produced something, which is probably to a large extent influenced by environmental factors. It also makes it unreliable for young persons, who may not have had time to develop their interests or potential. Clearly, children are unlikely to publish novels, attain patents, publish in academic journals, or whatever else might be regarded as evidence of creative achievement. It could also be averred that a person who produces art or other things that other people do not consider valuable or useful can nevertheless be creative, so we should be careful about using success as a measure of creativity. Someone can be highly creative, yet unrecognized, something true of many artists and writers who only become prominent posthumously.

Divergent thinking tests are poor measures of creativity, according to many scholars. Piffer (2012) reviewed these problems, offering the following arguments. (1) Creativity tests measure divergent thinking, but convergent thinking is also salient to creativity and that being the case, creativity may as well be measured with an IQ test. (2) Creativity can substantially be reduced to the personality trait Openness-Intellect combined with moderately low Agreeableness and moderately low Conscientiousness. It is also predicted by weak manifestations of a

* Corresponding author.
E-mail addresses: bakhiet@ksu.edu.sa (S.F.A. Bakhiet), guy.madison@umu.se (G. Madison), halmoghyrah@ksu.edu.sa (H.A.S. Almoghyrah), sabokhaled@ksu.edu.sa (S.H.S. Al-Mutairy), ysayed@ksu.edu.sa (Y.A.S. Essa).

https://doi.org/10.1016/j.actpsy.2022.103797
Received 5 March 2021; Received in revised form 27 October 2022; Accepted 17 November 2022
Available online 22 November 2022

number of pathologies such as schizotypal personality, hypomania and bipolar disorder (Piffer, 2018), which can all be tested by personality tests or other established instruments. (3) Creativity tests can at best be seen to measure 'creative potential,' as they do not inquire about lifetime creative achievement. (4) No general factor of creativity has so far been found, speaking against its construct validity. It can be responded, however, that the study of creativity is relatively new and developing, so we cannot expect it yet to be at the same theoretical level as that of intelligence. Further, some studies have found factors underlying the Torrance test, illustrating the utility of the creativity test which we have employed, if not of others (Fernando et al., 2007). The entire concept of Openness, and its conceptual validity, has been strongly disputed (Dutton & Charlton, 2015; Eysenck, 1995) and, anyway, creativity is likely to be better measured via external rather than self-assessment, with the latter being generally the case in personality tests. There is a degree to which originality, and thus divergent thinking, is more central to creativity than is convergent thinking, because it is divergent thinking that is associated with artistic accomplishment, while convergent thinking is more associated with scientific originality (Simonton, 2009). The artist is generally understood to be the most 'creative' as the term is usually defined (Nettle, 2007). Finally, even if it can be agreed that tests at best measure creative potential, they permit us to obviate the aspects of creative achievement which do not relate to creativity, such as knowing the right people, or having financial backing, and are, again, useful for assessing creativity amongst children.

Intelligence is also a component of creativity. Jensen (1998, p. 75) noted that creativity, specifically the divergent thinking component, is a robust correlate of general intelligence, with Hauck and Thomas (1972) finding a relationship of 0.29 between the two. Kim (2005) has found that, overall, the relationship between creativity and IQ is even weaker, at about 0.1. Logical, abstract and numerical reasoning have all been shown to correlate with the ability to produce original metaphor at around 0.4 (De Cassia Nakano et al., 2015). Eysenck (1995) has shown that creativity is a function of a combination of a moderate level of psychoticism (low Agreeableness and low Conscientiousness) combined with intelligence. This means that a relatively high IQ is required for creative accomplishments and, thus, creative accomplishments tend to increase as IQ increases. However, intelligence ceases to predict creative accomplishment above a certain point, although the precise breakpoint is unknown. Jauk et al. (2013) found that several different IQ breakpoints are reported in the literature, but that it is often unclear where these exact threshold values come from. Their own empirical estimation lands at IQ scores of 85 for fluency, 104 for originality of the top two responses, and 119 for average originality. The latter is similar to Eysenck's (1995) observation that IQ predicted creativity up to approximately 120 points, above which creativity was exclusively predicted by personality. However, Guilford (1950) argued that creative people may score lower than controls on IQ tests due to their approaching problems in unorthodox ways.

Scientific geniuses are generally defined as highly creative scientists who have made an extremely original and important contribution to their field (Simonton, 2009). Studies of such individuals concur that they tend to be characterized by a combination of very high IQ and moderate psychoticism (e.g. Dutton & Charlton, 2015; Eysenck, 1995; Post, 1994; Simonton, 2009). In addition, more quotidian traits have been found to be associated with creativity. For example, sense of humor is regarded as a distinct and socially influential trait, and it is associated with both creativity and intelligence (Howrigan & MacDonald, 2008; Miller & Tal, 2007). Evidently, then, creativity is important in predicting achievement of various kinds, but it is only marginally associated with IQ.

Torrance, who developed the creativity testing instrument upon which we will draw in this study, found that intelligent children who are troublesome at school or at home, because they are unusually full of ideas and questions, tend to be successful in creative fields later in life (Torrance, 1965). He also found that creativity involved intellectual risk-taking and ways of thinking that were divergent from the societal norm (Torrance, 1963). A particular merit of Torrance's test is that, in addition to measuring divergent thinking, it predicts creative achievement – in longitudinal studies - to a greater extent than do IQ tests (Kim, 2007, p. 117). This is important because it has been proposed that longitudinal studies are vital with regard to creativity, as some children undergo periods of 'slump' in creativity (Barbot et al., 2016). Accordingly, if people were becoming less creative over time this would potentially be a matter of concern due to the positive outcomes associated with creativity, but as shown above, it might also be associated with declining general intelligence. Kim (2011) found precisely this phenomenon in an American sample which she and her team re-analysed. Kim drew upon studies using the Torrance Test of Creative Thinking, which was developed in 1966, and which had been administered to large samples of people – ranging from kindergarten pupils through to 12th grade students and adults - in 1974, 1984, 1990, 1998, and 2008. This meant that she could compare average levels across time with a very large sample of 272,599 people, and she found that they had significantly decreased since 1990.

A number of researchers have opined on the reasons for these findings, with Dutton and Woodley of Menie (2018), for example, arguing that they parallel similar declines in many of the other measures of intelligence in Western countries, including IQ scores since around 1997, spatial perception and backward digit span. Indeed, since the late 1990s, average IQ scores have been falling in European countries (Dutton et al., 2016). It is very difficult to determine whether paper-and-pencil tests actually reflect change in 'real' phenotypic or genotypic intelligence, because they are indirect measures and partly influenced by test-taking ability. While such tests have excellent psychometric properties when assessing individual differences within otherwise comparable individuals, they cannot be assumed to be invariant across different groups. Societies differ in the level and content of education, as well as in cultural factors that might affect test-taking performance, and the same applies to differences within the same society across time. Therefore, we must look also to other indicators than paper-and-pencil tests. In this regard, the percentage of the Icelandic population carrying polymorphisms associated with very high educational attainment, and thus, indirectly, with high IQ, has also been falling in parallel with IQ (Kong et al., 2017).

Another powerful supporting observation is that simple reaction time, which is negatively correlated with IQ, has been increasing across generations (Woodley et al., 2014). This has even been shown for reaction times measured at the same time for people born in different years across merely 27 years in Sweden (Madison et al., 2016). However, heretofore nobody has explored this apparent decline in a developing country, wherein a variety of other factors that might relate to creativity – such as educational standards and IQ score (no matter how well this may or may not measure intelligence in such societies, see Dutton, Becker, et al., 2018 for a critique) – are going to be significantly different (Lynn & Becker, 2019). Indeed, studies of Western countries have shown that the nature of the learning environment has a significant impact on child creativity (Besançon & Lubart, 2008) through impacting how children are stimulated and how they cognitively develop (Lubart & Georgsdottir, 2004).

The present study contributes to this field by exploring possible changes across time in divergent thinking scores intended to assess creativity. Specifically, we compare divergent thinking in the Republic of Sudan, focusing on scores on fluency, flexibility and originality. A 2018 study demonstrated an increase in IQ scores in Khartoum between 2004 and 2016 (Dutton, Bakhiet, et al., 2018). This is a so-called Flynn- or Lynn-Flynn Effect (Lynn, 2019), referring to increasing scores on IQ tests, which is mainly found in Western countries during the twentieth century. There is evidence that these increases are not on *g* but rather on specialised abilities that weakly correlate with *g*. This means that the skills that underlie the solving of IQ test items have been pushed to their phenotypic maximum, as industrialised society has made people increasingly educated and increasingly think in a more scientific way (Flynn, 2012; see Egeland, 2022, for a review).

In that creativity is associated with IQ, we would predict that divergent thinking scores would have increased in Sudan across the same period. However, there are a number of factors that would militate against this. For example, Dutton, Bakhiet, Ziada, et al. (2017) reported a Negative Flynn Effect in Khartoum between 1999 and 2010, seemingly brought about by certain demographic changes in the Khartoum population (see Dutton, Bakhiet, et al., 2018). These changes included of widening of educational participation between 1999 and 2010 in Khartoum. The direction of a possible change in divergent thinking across time is therefore an open question.

2. Method

2.1. Participants

The Standardized Circles test from the Torrance battery for creative thinking was administered to children aged 8–12 years at the Al Qabas Schools. Al Qabas is a franchise of private schools in the cities of Khartoum, Omdurman, and Bahri. The 2005 sample consisted of 2130 pupils, 1075 boys and 1055 girls, and these data were collected in the context of a previous study (Bakhiet, 2006). We then administered the same tests in the same schools in the beginning of 2018 for ages 6–13, but used only the age groups that overlapped with the 2018 sample, which comprised 3922 pupils, 1874 boys and 2044 girls, as detailed in Table 1.

2.2. The Standardized Circles test

This is a subtest from the Torrance battery for creative thinking form B (Torrance, 1966). The test consists of 40 circles. The subject is given 10 min to draw the largest number of shapes and images using these circles as a primary part of each drawing they produce. The full instruction was:

Try to draw the largest number of images using the circles at the bottom of this page and the next page. Circles must be the primary part of each image. Then add pencil lines to the circles to complete the picture, so that you place marks inside or outside the circles, inside and outside together wherever you want. To draw the picture, try to think of things that no one has thought of, draw as many pictures or different topics as possible, and put as many ideas as possible in each picture. Then make these images tell a whole interesting story, and add a name for each image below it.[1]

The test yields three measures, representing the dimensions fluency, flexibility, and originality. Each dimension has a different evaluation method. Fluency is determined by adding all responses minus duplicate or non-related responses. Flexibility is calculated by combining the number of categories in which the responses are made. In determining the category, the rater should take into account the number of categories of responses that can be categorized as ‘human, household items, flowers, school tools, celestial bodies etc.’ Originality is based on the scarcity of the response, and the scarcity here is measured in relation to the actual responses that emerged from the performance of the 2005 sample. A response that is given by 4 % or more of the participants is assigned an originality score of one. Likewise, responses repeated by 3 to 3.99 % is assigned a score of two, 2.0 to 2.99 % the score three, 1 to 1.99 % a score of four, and responses that are repeated less than 1 % are assigned a score of five. These percentages are determined by the actual performance of the 2005 study sample. However, the distributions of these percentages were highly similar across the two years, when determined separately for each year.

[1] This is back-translated from Arabic, indicating its consistence with the original English instructions.

2.3. Procedure

All aspects of the data collection were highly similar across the two administrations of the test. This test consists of two pages. The first includes basic information about the pupil, and instructions for applying the test, while the second page includes the test consisting of (40) circles, which the pupil will answer in the specified time of ten minutes. According to the test administration procedures specified in the test manual of the Arabic version, the test implementer makes sure that each pupil writes his basic information, then the applicant begins to read the application instructions and asks the pupils to follow him while reading these instructions, taking into account that the pupil does not turn the page unless he is asked to do so. The administrator answers any inquiry, then asks the students to turn the page and each of them read the instructions shown at the head of the second page, and the test administer begins calculating the time allowed to answer, which is only 10 min. At the end of the first page there are some rectangles with symbols inside, and these rectangles are left to the correctors and the pupils do not write anything inside.

In both administrations (2005 and 2018), the administration procedures were in accordance with the instructions of the Arabic version of the test. In the 2005 administration, the main researcher and his assistants (holders of a Bachelor of Psychology trained in applying the test, correcting it, recording scores, and entering them into the computer) administered the test. In the first lesson, at the beginning of the school day, where the pupils are at the peak of their activity and their physical and mental readiness, the class teacher was hired to help maintain order and help the administration team in distributing test papers, while the main researcher provides instructions and explains the directions that precede the application, which are not limited in time, as this depends on the circumstances of each class.

3. Results

First, we consider the scores on Fluency, Flexibility, and Originality separately. Table 1 reports the descriptive data across age and sex, but separately for year of data collection (2005 or 2018). It shows that the skewness is within ±2.0 and the kurtosis is within ±7.0, which are considered acceptable limits for approximate normal distribution (George & Mallery, 2010).

The invariance across 2005 and 2018 is assessed in two ways, explorative factor analysis and zero-order correlations. The three first value columns of Table 2 shows the correlations, and the fourth column the factor loadings of principal component extraction factor analysis, separately for each year.

Figs. 1 through 3 plot the three indices as a function of age, sex, and year of data collection, separately for each year and across all ages and participants.

In general, the 2018 cohort is significantly less creative than the 2005 cohort. The Fluency scores were subjected to a three-way between-participants ANOVA, indicating significant main effects of year ($F_{1,\ 6028} = 31.49$, $p < .000001$, $\eta^2 = 0.0052$), sex ($F_{1,\ 6028} = 294.1$, $p < .000001$, $\eta^2 = 0.046$), and age ($F_{4,\ 6028} = 100.0$, $p < .000001$, $\eta^2 = 0.062$). All interactions were also significant, namely year x sex ($F_{1,\ 6028} = 45.75$, $p < .000001$, $\eta^2 = 0.0075$), year x age ($F_{4,\ 6028} = 9.449$, $p < .000001$), sex x age ($F_{4,\ 6028} = 6.22$, $p < .0001$), and year x sex x age ($F_{4,\ 6028} = 9.49$, $p < .000001$). Thus, females were very much more fluent in producing shapes and images than males, and fluency also increased with age, with the exception of males between 9 and 11 years of age in 2005. The main effect of year corresponds to higher fluency in 2005, but this was only true of the males, whereas the females did not differ between 2005 and 2018. Partial η^2 are reported for completeness, but more relevant for our questions are the main effect sizes of the year of administration. Cohen's d was computed as 0.120 across sex and age, and this difference was also significant according to an independent two-samples, two-tailed t-test, using the Welch-Satterthwaite equation (Welch, 1947) to compute the

Table 1
Descriptive data for each variable and year, across age and sex.

Variable	Year	N	Mean	SD	Min	Max	Skewness	Kurtosis
Fluency	2005	2130	10.14	4.94	1	41	1.078	2.747
	2018	3918	9.50	5.57	0	40	1.000	1.429
Flexibility	2005	2130	6.50	3.06	1	26	0.976	2.402
	2018	3918	5.89	3.32	0	33	1.205	3.441
Originality	2005	2130	6.96	7.098	0	54	1.682	4.150
	2018	3918	5.41	6.215	0	60	1.975	5.735
Composite	2005	2130	9.39	3.878	0.739	30.827	1.088	1.665
	2018	3918	7.68	4.85	0	42.14	1.160	2.315

Table 2
Zero-order correlations amongst the three sub-scores and their composite score, separately for each year and across all ages and participants, and factor loadings of the factor analysis (FI).

Year		Flexibility	Originality	Composite	FI
2005	Fluency	0.786	0.632	0.888	0.911
	Flexibility		0.634	0.860	0.910
	Originality			0.890	0.840
2018	Fluency	0.754	0.594	0.878	0.877
	Flexibility		0.715	0.895	0.926
	Originality			0.878	0.857

Note. Factor analysis explained variance is 78.87 % for 2005 and 78.79 % for 2018.

degrees of freedom ($t_{4829} = 4.62$, $p < .0001$). The means, standard deviations, and numbers of individuals are listed separately by sex, age, and year in Table A1 in the Appendix.

Fig. 2 depicts the Flexibility scores in the same fashion as in Fig. 1. The Flexibility scores were subjected to a three-way between-participants ANOVA, indicating significant main effects of year ($F_{1,\ 6028} = 58.78$, $p < .000001$, $\eta^2 = 0.0037$), sex ($F_{1,\ 6028} = 122.1$, $p < .000001$, $\eta^2 = 0.0076$), and age ($F_{4,\ 6028} = 120.7$, $p < .000001$, $\eta^2 = 0.029$). The effect of year was also significant across sex and age ($d = 0.187$, $t_{4679} = 7.13$, $p < .0001$). All second-order interactions were also significant, namely year x sex ($F_{1,\ 6028} = 4.58$, $p < .05$, $\eta^2 = 0.00029$), year x age ($F_{4,\ 6028} = 2.84$, $p < .05$), and sex x age ($F_{4,\ 6028} = 4.26$, $p < .005$), whereas the year x sex x age interaction was not ($F_{4,\ 6028} = 0.83$, $p = .50$). Again, females were very much more flexible in producing shapes and images from different categories than males, and this flexibility increased consistently with age for both sexes. The main effect of year corresponds to higher fluency in 2005, consistently across all ages and both sexes. Table A2 in the Appendix lists the means, standard deviations, and numbers of individuals separately by sex, age, and year.

The Originality scores are depicted in Fig. 3, which again shows that females perform better than males, that the performance increases with age for both sexes, and that it was higher in 2005 than in 2018. The Originality scores were subjected to a three-way between-participants ANOVA, indicating significant main effects of year ($F_{1,\ 6028} = 98.24$, $p < .000001$, $\eta^2 = 0.026$), sex ($F_{1,\ 6028} = 40.92$, $p < .000001$, $\eta^2 = 0.011$), and age ($F_{4,\ 6028} = 106.71$, $p < .000001$, $\eta^2 = 0.103$). The effect of year was also significant across sex and age ($d = 0.237$, $t_{3905} = 8.45$, $p < .0001$). All higher-order interactions were also significant, namely year x sex ($F_{1,\ 6028} = 10.558$, $p < .005$, $\eta^2 = 0.0028$), year x age ($F_{4,\ 6028} = 3.092$, $p < .05$), sex x age ($F_{4,\ 6028} = 5.290$, $p < .0005$), and year x sex x age ($F_{4,\ 6028} = 3.806$, $p = .005$). Table A3 in the Appendix lists the means, standard deviations, and numbers of individuals separately by

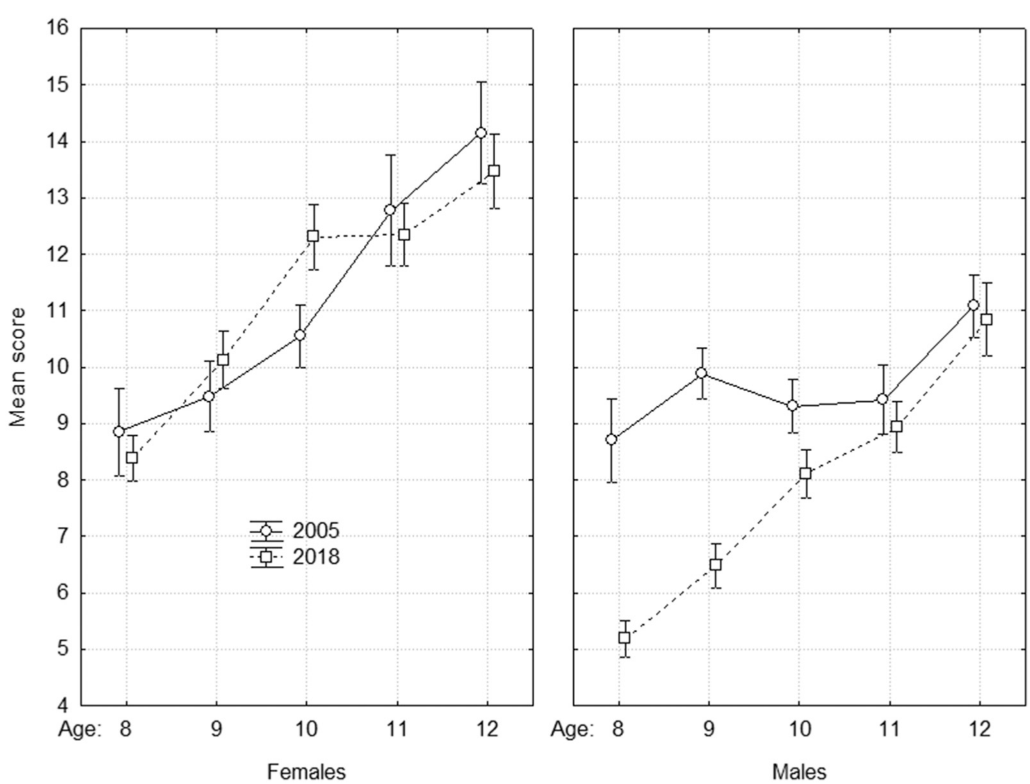

Fig. 1. Mean Fluency scores for year, age, and sex, across participants. Error bars depict 0.95 confidence intervals.

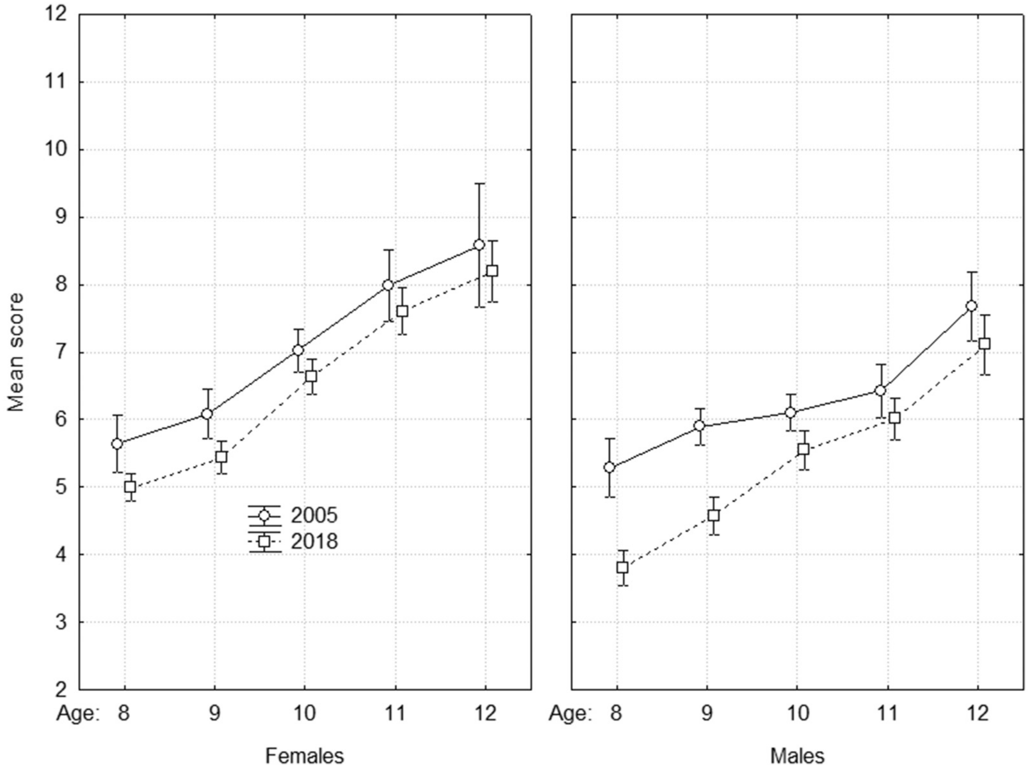

Fig. 2. Mean Flexibility scores for year, age, and sex, across participants. Error bars depict 0.95 confidence intervals.

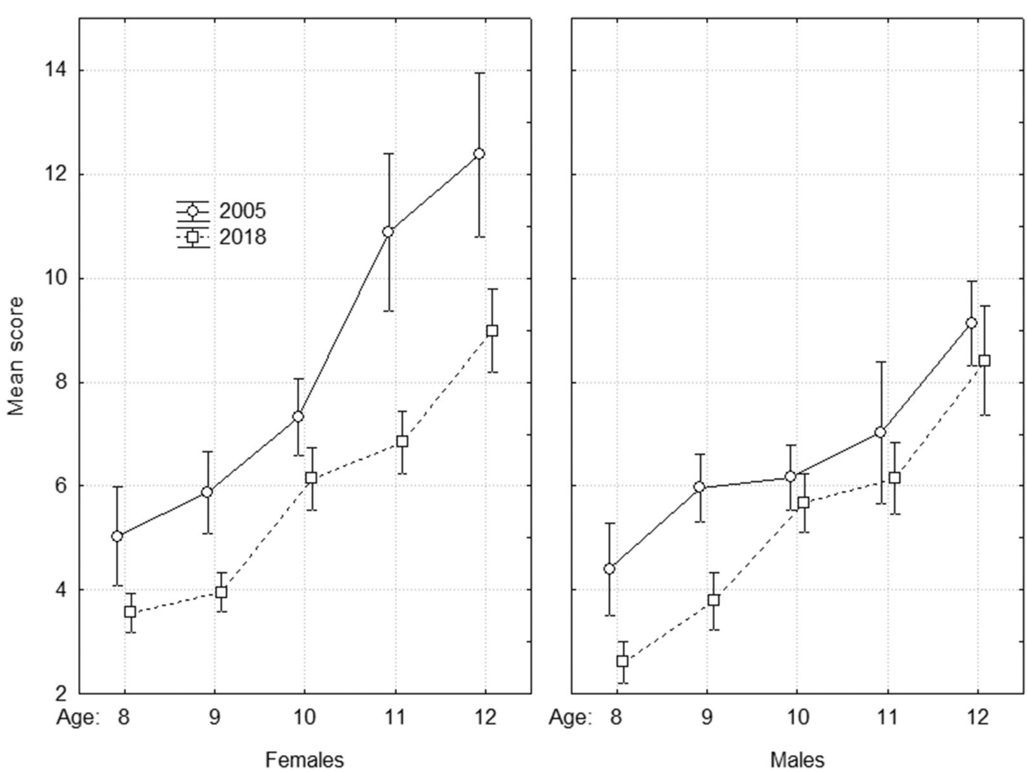

Fig. 3. Mean Originality scores for year, age, and sex, across participants. Error bars depict 0.95 confidence intervals.

sex, age, and year.

Finally, the mean of the three different indices of divergent thinking was taken to create a composite divergent thinking variable, depicted in Fig. 4. The distribution of these scores is slightly positively skewed and moderately leptokurtic, as plotted in Fig. A1 and detailed in Table 1.

They were subjected to a three-way between-participants ANOVA, indicating significant main effects of year ($F_{1,\ 6028} = 82.09$, $p < .000001$, $\eta^2 = 0.011$), sex ($F_{1,\ 6028} = 138.47$, $p < .000001$, $\eta^2 = 0.019$), and age

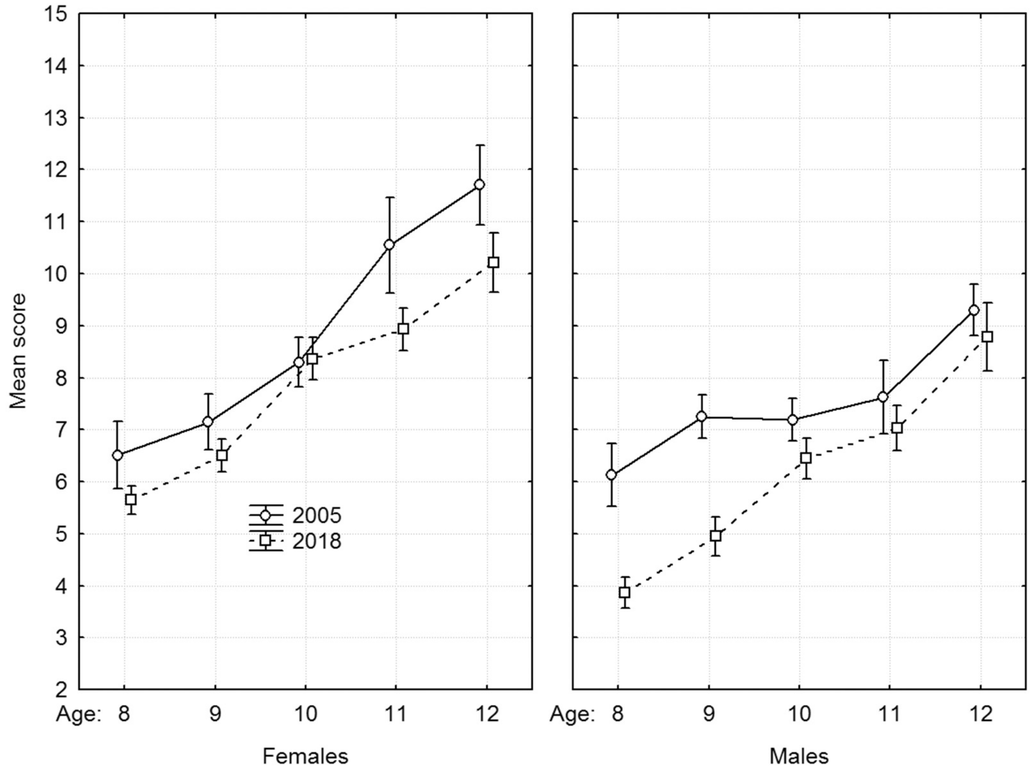

Fig. 4. Composite divergent thinking scores for year, age, and sex, across participants. Error bars depict 0.95 confidence intervals.

($F_{4,\ 6028} = 142.77$, $p < .000001$, $\eta^2 = 0.07$). The effect of year was also significant across sex and age ($d = 0.2097$, $t_{4355} = 7.74$, $p < .0001$). All higher-order interactions except year x sex ($F_{4,\ 6028} = 3.03$, $p = .081$) were significant, namely year x age ($F_{1,\ 6028} = 4.89$, $p < .001$, $\eta^2 = 0.0028$), sex x age ($F_{4,\ 6028} = 6.10$, $p < .0001$), and year x sex x age ($F_{4,\ 6028} = 5.52$, $p = .0005$). Table A4 in the Appendix lists the means, standard deviations, and numbers of individuals separately by sex, age, and year.

4. Discussion

The purpose of the present study was to examine if measures of divergent thinking have changed over time in the Republic of Sudan, in terms of fluency, flexibility, and originality. Our key finding is that divergent thinking scores have, indeed, significantly decreased in our sample between 2005 and 2018. In addition, scores have significantly decreased on each of the components. There are significant sex differences in the extent of this decrease. For example, the decline in originality was more pronounced amongst females while the fall in flexibility was more pronounced amongst males, for unknown reasons. As already noted, this decrease is not congruous with the finding by Dutton, Bakhiet, et al. (2018) that IQ scores increased in Khartoum between the years 2004 and 2016.

A possible limitation of the present conclusions is that the demographic structure of pupils may have changed in these 13 years. The public education system in Sudan is tremendously underfunded and insufficiently resourced. Accordingly, wealthier parents are increasingly removing their children from the public education system and placing them in private schools, such as Al Qabas schools (Husain, 2018). This would result in a Negative Flynn Effect in Al Qabas schools, as their pupils become decreasingly elite, with socioeconomic status of origin robustly correlating with IQ (Jensen, 1998, p.75), as the heritability of IQ is approximately 0.8 (Lynn, 2011, p. 101). Also, there may be a brain drain from the countryside towards Khartoum across this period, adding to the fact that the present sample are from a variety of areas and not merely from Khartoum, and so more rural than was that drawn upon by Dutton, Bakhiet, et al. (2018) and Dutton, Becker, et al. (2018). Similarly, the finding by Dutton, Bakhiet, Ziada, et al. (2017), of a Negative Flynn Effect in Khartoum between 1999 and 2010, may have been due to the expansion of free education in Sudan across this period. Accordingly, these results cannot be seen to contradict our own; they are products of sampling issues. Thus, we can be reasonably state that we have a prima facie case that creativity is decreasing in Sudan.

We are not aware of any such exodus from the public education system in developed countries, such as in Western Europe, and it is generally the case that, though school standards vary, the standard of the public education system is far higher in developed countries (e.g. Laabas, 2017), meaning that it is more likely to push pupils to their phenotypic maximum IQ (see Flynn, 2012). This would potentially be the simplest explanation for the decline in divergent thinking scores in Khartoum State Al Qabas schools, which this study has revealed, even though there is only a weak correlation between creativity and IQ. A growing number of studies have found a Negative Flynn Effect in Western countries, as already noted (Dutton et al., 2016) and there is evidence that this is on *g* (Woodley of Menie & Dunkel, 2015). Indeed, they have found it to be so for primarily genetic reasons. Specifically, alleles associated with very high educational attainment and thus indirectly with IQ have become less prevalent in European populations over the last three generations (Woodley of Menie et al., 2018; Kong et al., 2017). This could be happening in Arab countries as well. But it behooves researchers to be very cautious in assuming that negative Flynn Effects in Arab countries reflect the same process, due to these nations' suboptimal educational environment and related sociological changes. Negative Flynn Effects in Arab countries have now been reported twice in Sudan and once in Kuwait (Dutton, Bakhiet, Essa, et al., 2017) and there has been a cessation of the Flynn Effect in Syria (Dutton, Essa, et al., 2018). In all instances, the most parsimonious explanation is related to highly localized factors, such as changes in the education system that rendered it less efficient or less scientifically-focused.

Thus, our study adds to the literature indicating that areas such as intelligence, and its correlates, do not seem to follow the same processes in the Arab world as they do in the West. Intelligence in Arab countries appears to peak at a younger age and also increase more slowly than is the case in Western countries (Bakhiet et al., 2018), and where there is a Negative Flynn Effect it seems to have environmental, rather than even partly genetic, causes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through Research Group no. RG-1438-007.

Appendix A

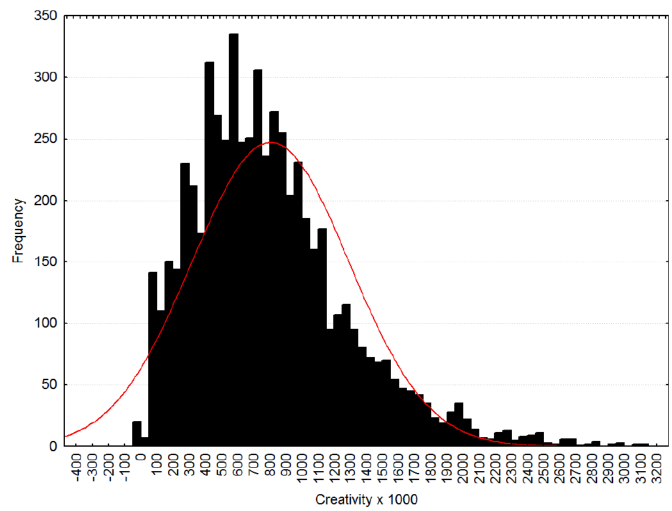

Fig. A1. Histogram of composite divergent thinking scores across all participants.

Table A1
Fluency raw scores for 2005 and 2018, for each age and sex.

Age	Sex	2005			2018		
		N	M	SD	N	M	SD
8	Males	120	8.70	4.05	431	5.18	3.34
	Females	150	8.85	4.77	438	8.38	4.20
9	Males	332	9.89	4.20	376	6.48	3.89
	Females	292	9.48	5.37	463	10.13	5.54
10	Males	368	9.30	4.58	445	8.11	4.56
	Females	354	10.55	5.29	415	12.31	5.97
11	Males	154	9.42	3.84	336	8.94	4.25
	Females	147	12.77	6.04	378	12.34	5.48
12	Males	101	11.08	2.83	286	10.85	5.55
	Females	112	14.15	4.75	350	13.47	6.21
Total males		1075	9.344	4.151	1874	7.914	4.320
Total females		1055	10.756	5.667	2044	10.328	5.482
All		2130	10.05	4.909	3918	9.121	4.901

Table A2
Flexibility raw scores for 2005 and 2018, for each age and sex.

Age	Sex	2005			2018		
		N	M	SD	N	M	SD
8	Males	120	5.28	2.38	431	3.81	2.70
	Females	150	5.64	2.63	438	4.99	2.14
9	Males	332	5.90	2.53	376	4.58	2.74

(continued on next page)

Table A2 (*continued*)

Age	Sex	2005			2018		
		N	M	SD	N	M	SD
	Females	292	6.08	3.20	463	5.44	2.64
10	Males	368	6.10	2.65	445	5.55	3.10
	Females	354	7.02	3.04	415	6.64	2.66
11	Males	154	6.43	2.46	336	6.01	2.84
	Females	147	7.98	3.24	378	7.61	3.48
12	Males	101	7.67	2.57	286	7.11	3.87
	Females	112	8.58	4.86	350	8.19	4.35
Total males		1075	5.95	2.566	1874	5.414	3.051
Total females		1055	7.06	3.439	2044	6.62	3.402
All		2130	6.505	3.003	3918	6.017	3.227

Table A3
Originality raw scores on Khartoum State Sample, 2005 and 2018.

Age	Sex	2005			2018		
		N	M	SD	N	M	SD
8	Males	120	4.40	4.972	431	2.61	4.11
	Females	150	5.03	5.927	438	3.56	3.96
9	Males	332	5.96	6.04	376	3.78	5.37
	Females	292	5.87	6.94	463	3.95	4.08
10	Males	368	6.16	6.08	445	5.67	6.11
	Females	354	7.32	7.08	415	6.15	6.20
11	Males	154	7.03	8.64	336	6.14	6.41
	Females	147	10.88	9.24	378	6.84	5.95
12	Males	101	9.14	4.11	286	8.41	9.10
	Females	112	12.37	8.40	350	9.00	7.62
Total males		1075	5.784	6.365	1874	4.466	5.222
Total females		1055	7.494	8.2	2044	5.244	5.466
All		2130	6.639	7.283	3918	4.855	5.344

Table A4
Composite divergent thinking scores on Khartoum State Sample, 2005 and 2018.

Age	Sex	2005			2018		
		N	M	SD	N	M	SD
8	Males	120	6.12	3.74	431	3.86	3.43
	Females	150	6.50	4.43	438	5.64	3.22
9	Males	332	7.24	4.24	376	4.94	3.99
	Females	292	7.15	5.02	463	6.50	3.69
10	Males	368	7.19	4.34	445	6.44	4.55
	Females	354	8.29	5.04	415	8.36	4.59
11	Males	154	7.62	4.89	336	7.03	4.39
	Females	147	10.54	6.15	378	8.92	4.48
12	Males	101	9.29	2.73	286	8.87	6.07
	Females	112	11.70	4.36	350	10.21	5.76
Total Males		1075	7.35	3.92	1874	6.01	4.40
Total Females		1055	8.39	4.91	2044	7.78	4.32
All		2130	8.93	5.33	3918	7.87	4.50

References

Bakhiet, S. (2006). Standardization of circles test from the Torrance battery for creative thinking form (B) on children aged 8–12 years at Al qabas School in Khartoum State. *Journal of Educational Studies, 14*, 102–137.

Bakhiet, S., Dutton, E., Ashaer, K., Essa, Y., Blahmar, T., Hakami, S., & Madison, G. (2018). Understanding the simber effect: Why is the age-dependent increase in children's cognitive ability smaller in Arab countries than in Britain? *Personality and Individual Differences, 122*, 38–42.

Barbot, B., Lubart, T. I., & Besançon, M. (2016). Peaks, slumps, and bumps: Individual differences in the development of creativity in children and adolescents. *New Directions for Child and Adolescent Development, 151*, 33–45.

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.

Woodley of Menie, M. A., Peñaherrera-Aguirre, M., Fernandes, H. B. F., & Figueredo, A. J. (2018). What causes the anti-Flynn effect? A data synthesis and analysis of predictors. *Evolutionary Behavioral Sciences, 12*, 276–295.

Woodley of Menie, M. A., & Dunkel, C. (2015). In France, are secular IQ losses biologically caused? A comment on Dutton and Lynn. *Intelligence, 53*(2015), 81–85.

De Cassia Nakano, T., Wechsler, S., Campos, C., & Millian, Q. (2015). Intelligence and creativity: Relationships and their implications for positive psychology. *Psico-USF, 20*, 2. https://doi.org/10.1590/1413-82712015200201

Dutton, E., Bakhiet, S., Essa, Y., Blahmar, T., & Hakami, S. (2017b). A negative Flynn effect in Kuwait: The same effect as in Europe but with seemingly different causes. *Personality and Individual Differences, 114*, 69–72.

Dutton, E., Bakhiet, S., Lynn, R., Essa, Y., Blahmar, T., Becker, D., Osman, H., & Hakami, S. (2018b). A Flynn effect in Khartoum, the sudanese capital: 2004–2016. *Intelligence, 68*, 82–86.

Dutton, E., Bakhiet, S., Ziada, K., Essa, Y., & Blahmar, T. (2017a). A negative Flynn effect in Khartoum, the sudanese capital. *Intelligence, 63*, 51–55.

Dutton, E., Becker, D., Osman, H., Bakhiet, S., Essa, Y., Ali, H., Alqafari, S., Hamdi, A., & Alfelah, A. (2018a). The Raven's test performance of south sudanese samples: A validation of criticisms of the utility of Raven's among sub-saharan africans. *Personality and Individual Differences, 128*, 122–126.

Dutton, E., & Charlton, B. (2015). *The genius famine*. Buckingham: University of Buckingham Press.

Dutton, E., & Woodley of Menie, M. A. (2018). *At Our Wits' End: Why We're Becoming More Intelligent and What It Means for the Future*. Exeter: Imprint Academic.

Dutton, E., Essa, Y., Bakhiet, S., Ali, H., Alqafari, S., Alfaleh, A., & Becker, D. (2018c). Brain drain in Syria's ancient capital: No Flynn effect in Damascus, 2004–2013/14. *Personality and Individual Differences, 126*, 10–13.

Dutton, E., Van der Linden, D., & Lynn, R. (2016). The negative Flynn effect: A systematic literature review. *Intelligence, 59*, 163–169.

Egeland, J. (2022). The ups and downs of intelligence: The co-occurrence model and its associated research program. *Intelligence, 92*. https://doi.org/10.1016/j.intell.2022.101643

Eysenck, H. J. (1995). *Genius: The natural history of creativity*. Cambridge: Cambridge University Press.

Fernando, M., Ferrándiz, C., Bermejo, M. R., Sánchez, C., Parra, J., & Prieto, M. D. (2007). Internal structure and standardised scores of the Torrance test of creative thinking. *Psicothema, 19*, 489–496.

Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the twenty first century*. Cambridge: Cambridge University Press.

George, D., & Mallery, M. (2010). *SPSS for windows step by step: A simple guide and inference, 17.0 update* (10th ed.). Boston, MA: Pearson.

Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444–454.

Hauck, W., & Thomas, J. W. (1972). The relationship of humor to intelligence, creativity, and intentional and incidental learning. *Journal of Experimental Education, 40*, 52–55.

Howrigan, D. P., & MacDonald, K. B. (2008). Humor as a mental fitness indicator. *Evolutionary Psychology, 6*, 652–666.

Husain, N. (2018). *Standardization of the standard progressive matrices plus (SPM+) in Khartoum State*. Sudan: Omdurman Islamic University.

Jauk, E. (2019). A bio-psycho-behavioral model of creativity. *Current Opinion in Behavioral Sciences, 27*, 1–6.

Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. *Intelligence, 41*, 212–221.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Kim, K.-H. (2005). Can only intelligent people be creative? A meta-analysis. *Journal of Secondary Gifted Education, 16*, 57–66.

Kim, K.-H. (2007). The two torrance creativity tests: The torrance test of creative thinking and thinking creatively in action and movement. In A.-G. Tan (Ed.), *Creativity: A handbook for teachers*. Singapore: World Scientific Publishing.

Kim, K.-H. (2011). The creativity crisis: The decrease in creative thinking scores on the Torrance tests of creative thinking. *Creativity Research Journal, 23*, 285–295.

Kong, A., Frigge, M., & Thorleifsson, G. (2017). Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Sciences, USA, 114*, E727–E732.

Laabas, B. (2017). *Arab development challenges of the new millennium*. London: Routledge.

Lubart, T. I., & Georgsdottir, A. (2004). Creativity: Development and cross-cultural issues. In S. Lau, A. N. N. Hui, & G. Y. C. Ng (Eds.), *Creativity: When east meets west* (pp. 23–54). Singapore: World Scientific Publishing.

Lynn, R. (2011). *Dysgenics: Genetic deterioration in modern populations*. London: Ulster Institute for Social Research.

Lynn, R. (2019). Reflections on sixty-eight years of research on race and intelligence. *Psych, 1*, 123–131.

Lynn, R., & Becker, D. (2019). *The intelligence of nations*. London: Ulster Institute for Social Research.

Madison, G., Woodley of Menie, M. A., & Sänger, J. (2016). Secular slowing auditory simple reaction time in Sweden (1959-1985). *Frontiers in Neuroscience, 10*, 407. https://doi.org/10.3389/fnhum.2016.00407

Miller, G. F., & Tal, I. R. (2007). Schizotypy versus intelligence and openness as predictors of creativity. *Schizophrenia Research, 93*, 317–324.

Nettle, D. (2007). *Personality: What makes you who you are*. Oxford: Oxford University Press.

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.

Piffer, D. (2018). The genetics of creativity: The underdog of behavior genetics? In R. E. Jung, & O. Vartanian (Eds.), *The Cambridge handbook of neuroscience of creativity*. Cambridge: Cambridge University Press.

Post, F. (1994). Creativity and psychopathology. *British Journal of Psychiatry, 165*, 22–34.

Sawyer, R. K. (2006). *Explaining creativity: The science of human innovation*. Oxford: Oxford University Press.

Simonton, D. K. (2009). Varieties of (scientific) creativity: A hierarchical model of domain-specific disposition, development, and achievement. *Perspectives on Psychological Science, 4*, 441–452.

Torrance, E. P. (1963). *Education and creative potential*. Minneapolis: University of Minnesota Press.

Torrance, E. P. (1965). *Gifted children in the classroom*. New York: Macmillan.

Torrance, E. P. (1966). *The Torrance tests of creative thinking– Norms: Technical manual research edition—Verbal tests, forms A and B— Figural tests, forms A and B*. Princeton, NJ: Personnel Press.

Welch, B. L. (1947). The generalization of "student's" problem when several different population variances are involved. *Biometrika, 34*, 28–35.

Woodley, M. A., te Nijenhuis, J., & Murphy, R. (2014). Is there a dysgenic secular trend towards slowing simple reaction time? Responding to a quartet of critical commentaries. *Intelligence, 46*, 131–147.