Personality and Individual Differences 88 (2016) 280–284

Contents lists available at ScienceDirect

Personality and Individual Differences

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

Showing their true colours: Possible secular declines and a Jensen effect on colour acuity — More evidence for the weaker variant of Spearman's Other Hypothesis

Michael A. Woodley of Menie [a,b,*], Heitor B.F. Fernandes [c]

[a] Department of Psychology, Technische Universität Chemnitz, Chemnitz, Germany
[b] Center Leo Apostel for Interdisciplinary Studies, Vrije Universiteit Brussel, Brussels, Belgium
[c] Department of Psychology, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

ARTICLE INFO

Article history:
Received 24 April 2015
Received in revised form 14 August 2015
Accepted 3 September 2015

Keywords:
100 Hue
Colour acuity
Co-occurrence model
Jensen effect
Secular decline
Spearman's Other Hypothesis

ABSTRACT

Spearman's Other Hypothesis predicts that the common factor amongst sensory discrimination measures corresponds to general intelligence (*g*). The co-occurrence model predicts that low-complexity physiological information-processing indicators reliably measure *g* across cohorts, and should therefore decline with time due to genetic changes in the broader population. As strong relations exist between general sensory discrimination and *g*, such measures should show evidence of secular declines. This is tested using *N*-weighted temporal regression of square-root Total Error Scores (√TES), obtained from four Western normative samples published in the 1980s, 90s and 2000s (combined $N = 752$) evaluated using the Farnsworth–Munsell 100-Hue colour acuity test (disattenuated *g* loading = .78). A significant temporal β value of .37 was found (controlling for national IQ), suggesting a decline in colour acuity equating to a reduction in *g* of −3.15 points per decade. Analysis of the subset of the cohorts aged 20–29 years, in which colour acuity is maximized, reveals a larger secular decline ($\beta = .67$, $N = 199$, −5.85 points per decade). The small number of studies employed in these analyses makes these findings tentative however. Also consistent with a weaker variant of the Other Hypothesis is the finding that 100-Hue acuity-IQ correlations are associated with the Jensen effect. The aggregate vector correlation across two studies is .63 ($N = 932.5$, $p < .05$).

1. Introduction

Sir Francis Galton (1883) was the first to propose that fine sensory discrimination might be associated with higher general intelligence (i.e. the cognitive processes common to the solving of diverse tests of mental ability; Jensen, 1998; Spearman, 1904), however the claim was not investigated empirically until Spearman (1904) conducted an analysis in which student performance on various sensory discrimination tasks (lightness, sound and weight) was found to correlate with teacher-ratings of their general intelligence (*g*). Spearman further posited the existence of a *general factor of discriminative ability* existing in the common factor variance amongst various sensory discrimination tasks. He argued that this should correlate nearly perfectly with *g*.

After many decades of neglect, in the 1990s and 2000s a series of papers by Ian Deary revisited what came to be termed *Spearman's Other Hypothesis* (Deary, 1994; Deary, 2000a, 2000b). The first direct test of the Other Hypothesis was conducted in 2004, when Deary and co-workers collected data on various sensory discrimination tasks amongst a sample of 62 Scottish secondary school students, along with various measures of IQ. Utilizing structural equations modelling (SEM) to estimate the common factor variance amongst the sensory discrimination and the cognitive ability measures, the latent general discrimination and *g* factors were found to correlate at .92, making them virtually isomorphic — consistent with the prediction of the Other Hypothesis. In a second analysis, Deary, Bell, Bell, Campbell, and Fazal (2004) reanalysed a much larger dataset (899 individuals) for which measures of both cognitive and sensory discrimination ability had been collected and analysed in a previous publication (Acton & Schroeder, 2001). Using the same SEM-based method it was found that *g* correlated with the general discrimination factor at .676 for the male and .681 for the female cohort, which indicated some divergence between the two common factors, but also demonstrated considerable shared variance, consistent with a weaker form of the Other Hypothesis.

Subsequent studies have demonstrated convergence between sensory discrimination and *g* using different modalities of sensory discrimination (e.g. Meyer, Hagmann-von Arx, Lemola, & Grob, 2010).

1.1. The co-occurrence model

IQ is a highly heterogeneous measure. Applying the basic bi-factor variance components model first proposed by Spearman (1904), IQ is

* Corresponding author at: Department of Psychology, Technische Universität Chemnitz, Chemnitz, Germany.
E-mail address: Michael.Woodley@vub.ac.be (M.A. Woodley of Menie).

http://dx.doi.org/10.1016/j.paid.2015.09.009
0191-8869/

broadly composed of general and specialized ability variances (*g* and *s*). As was mentioned previously, *g* variance relates to the cognitive processes common to the solving of diverse tests of mental ability whereas *s* variances relate to uncorrelated and specialized sets of processes that are specific to the domain of each ability.

It has been found that the magnitude of the negative correlation between IQ and fertility relates positively to the *g* loading of the test on which it is established, making it a *Jensen effect* (Woodley & Meisenberg, 2013). The *g* loadings of tests also positively relate to the strength of their association with phenotypic indicators of mutation load (i.e. the complement of deleterious mutations), such as measures of fluctuating asymmetry and factors that increase mutation load, such as inbreeding depression (Prokosch, Yeo, & Miller, 2005; Rushton & Jensen, 2010). IQ subtests that are more *g*-loaded are also more heritable on average (see: Woodley of Menie, Fernandes & Hopkins, 2015 for a review of this literature), indicating that genetic factors influence individual differences in *g* to a greater extent than *s*.

At the opposite end of the continuum are variables that influence IQ primarily via environmental routes, such as educational interventions (te Nijenhuis, Jongeneel-Grimen & Kirkegaard, 2014), IQ gains amongst adopted children (te Nijenhuis, Jongeneel-Grimen & Armstrong, 2015), IQ gains via retesting effects (te Nijenhuis, van Vianen & van der Flier, 2007), and the Flynn effect (te Nijenhuis & van der Flier, 2013). In all cases, the effects of these environmental factors are bigger when the sub-scales are *g*-loaded to a lesser (and *s* loaded to a correspondingly greater) extent.

This observation has led to the development of the *co-occurrence model* which posits that genetic changes occurring within modern populations should be reducing the level of *g*, whereas environmental improvements of one sort or another should be increasing performance with respect to numerous *s*-variances associated with cognitive abilities simultaneously.

Whilst the gain in *s* is well-attested on the basis of ubiquitously rising IQ scores, there are several lines of evidence supporting co-occurrent declines in indicators of *g*, such as historical declines in per capita macro-innovation and genius, suggesting falling creativity (Huebner, 2005, Murray, 2003), diminishing backwards digit span (a working memory measure; Woodley of Menie & Fernandes, 2015), and decreased usage of high-difficulty words over 155 years, across a representative corpus of the English language, indicating declining crystallized ability, specifically vocabulary in direct response to genetic selection (Woodley of Menie, Fernandes, Figueredo & Meisenberg, 2015).

More germane to the issue of sensory discrimination is the finding of an apparent decline in performance on measures of simple reaction time (SRT) since the 19th century (Silverman, 2010; Woodley, te Nijenhuis & Murphy, 2013). This finding was not received uncritically (e.g. Dodonova & Dodonov, 2013). Reanalysis of a more closely matched subset of the original data to which various adjustments for methods variance had been made revealed declines of between −.57 and −1.21 points per decade, based on different applications of the corrections (Woodley, te Nijenhuis & Murphy, 2014). Tentative indications of generational slowing of SRT have also been found by comparing the extrapolated longitudinal to the observed cross-sectional ageing trend amongst various studies (Verhaeghen, 2014). A more detailed analysis of three Scottish birth cohorts utilizing a variant of this method, found that amongst the female cohorts the secular slowing in SRT performance equates to a decrease in *g* of −1.8 points per decade (Woodley, Madison & Charlton, 2014).

Even though SRT does not correlate strongly with IQ (−.31, rising to −.54 when corrected for range restriction, reliability and validity; Woodley et al., 2013), it is nonetheless likely a stable measure of a phenotype fundamental to *g* (i.e. information processing speed) across cohorts, as its simplicity reduces its sensitivity to training effects (Jensen, 2006), which may actually be increasing performance on more complex measures of processing speed (Verhaeghen, 2014). Pencil and paper tests on the other hand are sensitive to training effects and changes in test taking habits, therefore they frequently fail to measure the same parameter across cohorts (Wicherts et al. 2004). This lack of measurement invariance across cohorts is the reason why tests can sometimes be highly *g*-loaded within cohorts (such as the Raven's Progressive Matrices) but may nonetheless yield large Flynn effects between cohorts, tracking instead the development of specialized skills and abilities (e.g. Fox & Mitchum, 2013).

Sensory discrimination tasks are therefore potentially ideal for tracking secular trends in *g* over time, as like SRT, they exhibit low-complexity and are potentially resistant to training, therefore they should be relatively stable measures of a fundamental psychophysical indicator of *g* across generations.

1.2. The present study

Here the focus will be on colour discrimination ability evaluated using the Farnsworth–Munsell 100-Hue colour perception test (Farnsworth, 1943), which was first found to correlate with IQ in the 1960s (Lakowski, 1970). This test evaluates colour acuity by having the participants physically arrange a series of 85 caps, each of subtly different hue, along a spectrum defined by two end caps (e.g. blue to green, pink to purple etc.). Normative performance data have been collected from four Western populations between the 1980s and 2000s, which will here be reanalysed for the presence of possible secular trends.

In order to validate potential declines in colour acuity with respect to potential underlying declines in *g*, an analysis involving the method of correlated vectors will be performed by reanalysing data from Acton and Schroeder (2001) and Deary and co-workers (2004), which provide information on subtest-100-Hue acuity correlations, along with *g* loadings. This is an important relationship to establish as single indicators can contribute to a *g*-loaded aggregate even when, taken on their own, they relate more to specialized abilities (e.g. Inspection Time; Deary & Crawford, 1998).

2. Methods

2.1. Secular trend estimation

Secular trend data on 100-Hue acuity were obtained using four normative studies published in the 1980's (Verriest, Van Laetham, & Uvijls, 1982), 90s (Roy, Podgor, Collier, & Gunkel, 1991), and 2000s (Kinnear & Sahraie, 2002; Mäntyjärvi, 2001). These studies were conducted on mixed-sex samples of 232 Belgian, 112 American, 160 Finnish and 286 British subjects with the greatest age range extending from five to >80 years. Each of these studies employed the physical 'hardware' variant (as opposed to a more recent electronic variant) of the 100-Hue test. Very few studies have attempted to examine the normative characteristics of the 100-Hue test. Earlier normalizations were carried out on a sample of Americans in the 1950s (Farnsworth, 1957) and also amongst a Belgian sample in the 1960s (Verriest, Vandevyvere, & Vanderdonck, 1962), however owing to differences in the instrumentation used and also the statistical procedures employed in analysing the data, only the results of the more recent four studies are directly comparable to one another at the methodological level (Kinnear & Sahraie, 2002). There also exists biocultural heterogeneity between the countries, which potentially restricts comparability. A major source of this is national IQ, which ranges from 97.5 (USA) to 100.9 (Finland) (Lynn & Vanhanen, 2012).

100-Hue test performance is not associated with sex differences (Kinnear & Sahraie, 2002), however it is strongly age-dependent, with the highest Total Error Scores (the sum of the cap arrangement errors — these values are typically transformed by taking the square-root in order to normalize their distributions [$\sqrt{TES}$]) being found amongst very young samples (i.e. <10 years of age) and those in old age. The lowest $\sqrt{TES}$ values are typically found amongst young adults in their 20s. In their comparative analysis of three normative samples, Kinnear and Sahraie (2002) compared the different age groups in each study with one another using bar charts (p. 1410). The comparisons involving the groups aged 15–19, 20s, 60s and 70s all reveal patterns suggestive of a monotonic secular trend towards worsening colour acuity. Here the overall acuity of the four samples will be compared by computing √TES values using three different methods (mean, median and midrange) across the age groups. These will then be combined into a multi-method sample-level mean. Data on those < 10 years, and > 69 years of age will be excluded as not all studies report data for these age cohorts. Thus all values will be computed with respect to precisely age-matched cohorts in all four cases (10–69 years). The mean √TES values for those aged in their 20s (i.e. the proportion of the samples that have attained maximum acuity) will also be analysed in order determine whether the secular trend is robust to controlling for age.

Methods variance exists between those studies that evaluated colour acuity on a binocular (i.e. both eyes) rather than on a monocular (i.e. one eye at a time) basis. Verriest et al. (1982) administered the 100-Hue test to the same participants under both conditions. The average √TES value across the monocular right and left eye conditions is .93 units greater than under the binocular condition, indicating that the monocular condition reduces accuracy. Similarly the pooled binocular standard deviation value is .55 units lower than the monocular one, indicating higher intra-individual variance and therefore lower accuracy. To control for this, each right-left eye mean √TES value for each age group in Roy et al. (1991) and Mäntyjärvi (2001) is corrected downwards by .93 units. Each standard deviation value is similarly corrected downwards by .55 units. Descriptive statistics for √TES are presented in Table 1.

The secular trends will be estimated utilizing *N*-weighted least squares temporal regression implemented in SAS v.9.0 with a General Linear Model (GLM) function, using the study year (data collection years were not reported) as the independent variable and controlling for national IQ differences hierarchically, that is, IQ will be entered into the regression first (national IQ values are sourced from Lynn & Vanhanen, 2012). Significance and 95% Confidence Intervals will be computed meta-analytically, using the combined *N* of participants (752 in the case of the full-sample and 119 in the case of the 20–29 year old samples). Weighting by *N*, as opposed to standard error of the mean or inverse of the variance is appropriate in this case as the √TES values are scaled equivalently across studies. This simulates the results that would be obtained in performing a secondary analysis of the pooled raw data.

2.2. Computing IQ changes

Based on the results of *N*-weighted regression, secular changes in *g* will be calculated using the following steps (based on Woodley et al., 2013):

I) √TES values will be computed for 1982 and 2002 using the regression formula.

II) The change in √TES will be rescaled in terms of standard-deviation units via division by standard deviation. This parameter is calculated by pooling the age-specific standard deviation values for each sample, and then pooling these, yielding a value of 1.98. For the 20–29 year old cohort, the standard deviations are pooled across studies, yielding a value of 1.69.

III) The change in standard deviation units will be rescaled in terms of the change in *g* by division by the *g* loading of the 100-Hue test. The *g* loading of this test can be estimated using the results of the three SEMs in Deary et al. (2004). Taking the loadings of the general sensory discrimination factor on the 100-Hue test and then dividing these by the path coefficient between this factor and *g* (correcting for psychometric validity; Jensen, 1998, p. 383), yields an *N*-weighted disattenuated aggregated *g* loading of .78.

IV) The results of step III are then multiplied by 15 (the 'standard' IQ standard deviation) in order to rescale them in terms of IQ points, scaled as *g*, lost over 20 years. Division by two yields points lost per decade.

2.3. Method of correlated vectors

The method of correlated vectors will be utilized in order to determine whether the strength of the association between √TES and IQ relates to the *g* loading of the subtest on which the association is established. This is achieved by taking the Pearson correlation between the vector of subtest *g* loadings and the vector of the subtest-colour acuity correlation (*d*).

Two separate vector correlations will be computed; the first will employ subtest-colour acuity correlations from Acton and Schroeder (2001, p. 266), available on each of the 13 cognitive ability tests comprising the Johnson O'Connor Research Foundation Battery. The *g* loadings for these subtests are taken from the SEMs for males and females separately estimated by Deary and co-workers (2004, pp. 13–14). An unweighted average (representing a 50:50 sex ratio) of these is then computed for each subtest. The second analysis employs the three batteries administered to the sample of 62 Scottish secondary school children. The battery-acuity correlations are available from the correlation matrix (p. 7), and the *g* loadings are taken from the two-factor SEM (p. 9). The significances are computed by weighting the parameter estimates of the vectors being correlated by the sample sizes upon which they are based, as would normally be done in the context of meta-analytic data aggregation.

The resultant vector correlations will then be aggregated using software publically available at http://vassarstats.net.

3. Results

3.1. Secular trend

Fig. 1 presents the secular trends across the four normative samples for the total age-range (10–69, solid line) and for those aged 20–29 (dashed line).

Table 1
Descriptive statistics for the four normative samples reporting 100-Hue √TES values.

Study (all age ranges)	Country	Mean √TES value	Midrange √TES value	Median √TES value	Multi-method mean √TES value	Pooled standard deviation	Sample size (*N*)	National IQ
Verriest et al. (1982)	Belgium	7.77	7.63	8	7.8	2.32	205	99.3
Roy et al. (1991)[†]	USA	6.16	6.78	5.57	6.17	1.88	97	97.5
Mäntyjärvi (2001)[†]	Finland	7.85	7.87	8.07	7.93	1.78	154	100.9
Kinnear & Sahraie (2002)	UK	8.59	8.72	8.46	8.59	1.81	296	99.1
20-29 age range								
Verriest et al. (1982)	Belgium	5.69				2.07	29	
Roy et al. (1991)	USA	5.07				1.78	25	
Mäntyjärvi (2001)	Finland	6.57				1.84	30	
Kinnear & Sahraie (2002)	UK	6.73				1.1	35	

[†] Indicates that the study reports √TES values separately for each eye.

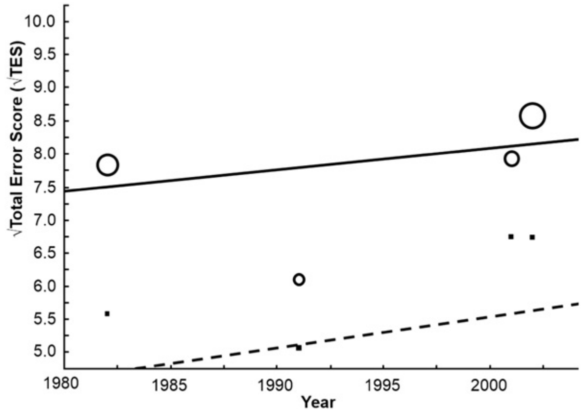

Fig. 1. Secular trends for multi-method √TES means (all age ranges, solid line; participant $N = 752$) and mean values for the samples aged 20–29 (dashed line, participant $N = 119$) based on the result of *N*-weighted temporal regression, controlling for population differences in IQ. Bubble size is proportional to sample size.

For the multi-method √TES means encompassing the full age range, the β value is .37, which is statistically significant given a participant *N* of 752 (95% Confidence Interval = .31 to .43). Based on the regression formula ($-58.718 + .033 * \text{year}$), the √TES value in 1982 was 7.5. The value in 2002 was 8.16, a difference of .66 √TES units. Dividing this by the cross-study pooled standard deviation value (1.98) yields a *d* value of .33. Dividing this by the *g* loading of the 100-Hue acuity test yields a *d* of *g* of .42. Multiplying by 15 yields a *g* decline of −6.3 points over 20 years, or −3.15 points per decade.

For the 20–29 year old subsamples, the β value is .67, which is statistically significant given a participant *N* of 119 (95% CI = .56 to .76). Based on the regression formula ($-96.397 + .051 * \text{year}$), the √TES value in 1982 was 4.68. The value in 2002 was 5.71, a difference of 1.03 √TES units. Dividing this by the cross-study pooled standard deviation value (1.69) yields a *d* of .61. Dividing this by the 100-Hue acuity *g* loading yields a *d* of *g* of .78. Multiplying by 15 yields a *g* decline of −11.7 points over 20 years, or −5.85 points per decade.

3.2. MCV analysis

The results of the MCV analyses are presented in Table 2.

4. Discussion

N-weighted temporal correlation involving multi-method mean √TES values computed across four normalization samples, matched by age range and controlled for national IQ suggests a secular increase in √TES values (decreasing colour acuity). The same analysis performed on the subset that has maximum colour acuity (aged 20–29) also reveals a potential secular increase. When rescaled in terms of the decline in *g* (scaled as IQ points), a loss of −2.22 points per decade is found when all age ranges are considered, increasing to −5.85 points per decade when just the samples aged 20–29 are considered.

Four observations may not seem like much on which to base a secular trend, however it is important to note that these four normative studies represent the virtual totality of the data that can be compared, as they are closely matched in terms of methodology, statistical treatment, participant recruitment procedure and also age. It must also be reiterated that the four studies are not sourced from the same countries, therefore whilst certain know sources of between-study heterogeneity associated with cognitive measures have been controlled, there may nonetheless exist other potential sources that have not been accounted for. It is on this basis that the secular trend must be considered tentative.

The observed declines are very large — much larger than that which might reasonably be expected on the basis of genetic selection (i.e. −.39 points per decade; Woodley of Menie, 2015). The declines, when the full age ranges are considered, are however consistent in magnitude with those observed in the UK on the Piagetian Volume and Heaviness task (−4.26 IQ points per decade; Shayer & Ginsburg, 2007), which hints at possible commonalities as both sets of tasks seem to engage discriminative faculties, i.e. noticing similarities across manipulated physical quantities in the case of Piagetian conservation tasks, and subtle differences in the case of sensory discrimination tasks. The declines are approximately twice as large when the 20–29 year old cohorts are considered. It is possible that the small sample sizes in the 20–29 year age-range cohort may be a source of sampling error, which is leading to an overestimation of the trend magnitude.

The magnitude of the decline could also be due to factors that are working synergistically with genetic changes, such as changing population composition via immigration (African and Middle Eastern populations exhibit higher √TES scores relative to US and European populations; Karaca, Saatçi & Kaynak, 2005; Wa Kaimbo & Missotten, 1994) or possibly neurotoxin exposure (Gobba & Cavalleri, 2003). Alternatively, the relatively small sample sizes within each age cohort suggest possible range restriction, which might be leading to a systematic

Table 2
g loadings and subtest/battery-acuity correlations (d) from two studies, along with the vector correlations computed for each separately and aggregated meta-analytically. Correlations are reverse scored so that greater acuity on the 100-Hue test corresponds to higher IQ.

Subtest	*g* loading	*d* (subtest-acuity)	Battery	*g* loading	*d* (battery-acuity)
Memory for design	.74	.28	Mill Hill vocabulary test	.55	.25
Number memory	.54	.16	Cattell culture fair test	.8	.45
Number series	.63	.29	Digit symbol test	.53	.28
Analytical reasoning	.63	.21			
Paper folding	.66	.32			
Observation	.57	.16			
Silograms	.50	.09			
Wiggly block	.62	.23			
Number facility	.55	.16			
English vocabulary	.24	.23			
Inductive reasoning	.50	.19			
Number checking	.33	.08			
Ideophoria	.25	.14			
$r (g * d)$	**.56*** ($N = 871.5$†)		$r (g * d)$	**.98*** ($N = 61$)	
$\rho (g * d)$	**.63*** ($N = 932.5$†)	95% CI = .59 to .66			

Note. *$p < .05$ †N based on averaging male and female participants from Deary et al.'s (2004) second analysis and those form Acton and Schroeder (2001).

underestimation of the 'true' population level √TES standard deviation value. If the true value is found to be larger, then when used as a reference estimate, the magnitude of the observed secular decline will necessarily decline, perhaps to the levels expected based on the existing data on SRT.

The finding of a possible secular decline and a large-magnitude Jensen effect on 100-Hue acuity would nonetheless seem to support the weaker variant of Spearman's Other Hypothesis, as both phenomena indicate affinities between sensory (in this case colour) discrimination and *g*.

References

Acton, G. S., & Schroeder, D. H. (2001). Sensory discrimination as related to general intelligence. *Intelligence*, *29*, 268–314.

Deary, I. J. (1994). Sensory discrimination and intelligence: Postmortem or resurrection? *American Journal of Psychology*, *107*, 95–115.

Deary, I. J. (2000a). *Looking down on human intelligence.* Oxford, UK: Oxford University Press.

Deary, I. J. (2000b). Simple information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence*. Cambridge, UK: Cambridge University Press.

Deary, I. J., Bell, J. P., Bell, A. J., Campbell, M. L., & Fazal, N. D. (2004). Sensory discrimination and intelligence: Testing spearman's other hypothesis. *American Journal of Psychology*, *117*, 1–18.

Deary, I. J., & Crawford, J. R. (1998). A triarchic theory of jensenism: Persistent conservative reductionism. *Intelligence*, *26*, 273–282.

Dodonova, Y. A., & Dodonov, Y. S. (2013). Is there any evidence of historical slowing of reaction time? No, unless we compare apples and oranges. *Intelligence*, *41*, 674–687.

Farnsworth, D. (1943). The Farnsworth-Munsell 100-hue and dichotomous tests for colour vision. *Journal of the Opthalmological Society of America*, *33*, 568–578.

Farnsworth, D. (1957). *The Farnsworth–Munsell 100-Hue test for the examination of color discrimination. Revised 1957.* Maryland: Munsell Color Company.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge based theory of rising scores on "culture-free" tests. *Journal of Experimental Psychology: General*, *142*, 979–1000.

Galton, F. (1883). *Inquiries into human faculty and its development.* London: Macmillan.

Gobba, F., & Cavalleri, A. (2003). Color vision impairment in workers exposed to neurotoxic chemicals. *Neurotoxicology*, *24*, 693–707.

Huebner, J. (2005). A possible declining trend for worldwide innovation. *Technological Forecasting and Social Change*, *72*, 980–986.

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

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences.* Amsterdam: Elsevier.

Karaca, A., Saatçi, A. O., & Kaynak, C. (2005). The result of the Farnsworth–Munsell 100 Hue test in Turkish population. *Ret-Vit*, *13*, 119–123.

Kinnear, P. R., & Sahraie, A. (2002). New Farnsworth–Munsell 100 hue test norms of normal observers for each year of age 5–22 and for age decades 30–70. *British Journal of Opthalmology*, *86*, 1408–1411.

Lakowski, R. (1970). Psychological variables in colour vision testing. In M. Rishter (Ed.), *Proceedings of the First International Colour Congress, Stockholm* (pp. 79–86). Gottingen: Muster-Schmidt.

Lynn, R., & Vanhanen, T. (2012). *Intelligence: A unifying construct for the social sciences.* London: Ulster Institute for Social Research.

Mäntyjärvi, M. (2001). Normal test scores in the Farnsworth–Munsell 100 Hue test. *Documenta Opthalmologica*, *102*, 73–80.

Meyer, C. S., Hagmann-von Arx, P., Lemola, S., & Grob, A. (2010). Correspondence between the general ability to discriminate sensory stimuli and general intelligence. *Journal of Individual Differences*, *31*, 46–56.

Murray, C. (2003). *Human accomplishment: The pursuit of excellence in the arts and sciences, 800 BC to 1950.* New York: Harper Collins.

Prokosch, M. D., Yeo, R. A., & Miller, G. F. (2005). Intelligence tests with higher *g* loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence*, *33*, 203–213.

Roy, M. S., Podgor, M. J., Collier, B., & Gunkel, R. D. (1991). Color vision and age in a normal North American population. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *229*, 139–144.

Rushton, J. P., & Jensen, A. R. (2010). The rise and fall of the Flynn effect as a reason to expect the narrowing of the black–white gap. *Intelligence*, *38*, 213–219.

Shayer, M., & Ginsburg, D. (2007). Thirty years on — A large anti-Flynn effect? The Piagetian test volume & heaviness norms 1975–2003. *Journal of Educational Psychology*, *77*, 25–41.

Silverman, I. W. (2010). Simple reaction time: It is not what it used to be. *American Journal of Psychology*, *123*, 39–50.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

te Nijenhuis, J., Jongeneel-Grimen, B., & Armstrong, E. L. (2015). Are adoption gains on the *g* factor? A meta-analysis. *Personality & Individual Differences*, *73*, 56–60.

te Nijenhuis, J., Jongeneel-Grimen, B., & Kirkegaard, E. O. W. (2014). Are Headstart gains on the *g* factor? A meta-analysis. *Intelligence*, *46*, 209–215.

te Nijenhuis, J., & van der Flier, H. (2013). Is the Flynn effect on g? A meta-analysis. *Intelligence*, *41*, 802–807.

te Nijenhuis, J., van Vianen, A. E. M., & van der Flier, H. (2007). Score gains on *g*-loaded tests: No *g*. *Intelligence*, *35*, 283–300.

Verriest, G., Vandevyvere, R., & Vanderdonck, R. (1962). Nouvelles recherches se rapportant à l'influence du sexe et de l'age sur la discrimination chromatique, ainsi qu'à la signification pratique des resultats du test 100 hue de Farnsworth–Munsell. *Revue d'Optique Théorique et Instrumentale*, *41*, 499–509.

Verhaeghen, P. (2014). *The elements of cognitive aging: Meta-analyses of age-related differences in processing speed and their consequences.* Oxford: Oxford University Press.

Verriest, G., Van Laetham, J., & Uvijls, A. (1982). A new assessment of the normal ranges of the Farnsworth–Munsell 100-Hue test scores. *American Journal of Ophthalmology*, *93*, 635–642.

Wicherts, J. M., Dolan, C. V., Hessen, D. J., Oosterveld, P., van Baal, G. C. M., Boomsma, D. I., et al. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, *32*, 509–537.

Wa Kaimbo, K., & Missotten, S. W. (1994). The Farnsworth–Munsell 100 Hue test in the Bantu population preliminary results. *Journal Français d'Opthalmologie*, *17*, 664–667.

Woodley, M. A., Madison, G., & Charlton, B. G. (2014). Possible dysgenic trends in simple visual reaction time performance in the Scottish twenty-07 cohort: A reanalysis of Deary & Der (2005). *Mankind Quarterly*, *55*, 110–124.

Woodley, M. A., & Meisenberg, G. (2013). A Jensen effect on dysgenic fertility: An analysis involving the National Longitudinal Survey of Youth. *Personality and Individual Differences*, *55*, 279–282.

Woodley, M. A., te Nijenhuis, J., & Murphy, R. (2013). Were the Victorians cleverer than us? A decline in general intelligence estimated from a meta-analysis of the slowing of simple reaction time. *Intelligence*, *41*, 843–850.

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.

Woodley of Menie, M. A. (2015). How fragile is our intellect? Estimating losses in general intelligence due to both selection and mutation accumulation. *Personality and Individual Differences*, *75*, 80–84.

Woodley of Menie, M. A., & Fernandes, H. B. F. (2015). Do opposing secular trends on backwards and forwards digit span evidence the co-occurrence model? A comment on gignac (2015). *Intelligence*, *50*, 125–130.

Woodley of Menie, M. A., Fernandes, H. B. F., Figueredo, A. J., & Meisenberg, G. (2015). By their words ye shall know them: evidence of genetic selection against general intelligence and concurrent environmental enrichment in vocabulary usage since the mid 19th century. *Frontiers in Psychology*, *6*, 361.

Woodley of Menie, M. A., Fernandes, H. B. F., & Hopkins, W. D. (2015). The more *g*-loaded, the more heritable, variable and evolvable: Homology with humans in chimpanzee cognitive abilities. *Intelligence*, *50*, 159–163.