Estimating One's Own and Other's Psychological Test Scores

Adrian Furnham

BI: Norwegian Business School, Oslo, Norway

Email: a.furnham@ucl.ac.uk

Abstract

This paper examines how accurate people are at estimating their own psychometric test results, which assess personality, intelligence, approach to learning and other factors. Seven groups of students completed a battery of power (general intelligence, fluid intelligence, creativity and general knowledge) tests and preference (approaches to learning, emotional intelligence, Big Five personality) tests. Two months later (before receiving feedback on their psychometric scores) they estimated their own scores and that of a class acquaintance who they claimed to know well on these variables. Results from the different samples were reasonably consistent. They showed that participants could significantly predict/estimate their own Neuroticism, Extraversion and Conscientiousness scores, as well as their General, Fluid and Crystalised intelligence, Approaches to Learning, Creativity and Happiness. Correlations between estimated and test-derived scores for an acquaintance were around half those for self-estimates and better for personality than ability. Participants self and "other" estimates were nearly all significantly positive. The discussion considers when, if ever, self-estimated scores can be used as proxy for test scores and what self-estimated scores indicate. Limitations are considered.

Keywords

Personality, Intelligence, Creativity, Learning Style, Self-Estimation, Self-Assessment

1. Introduction

There are psychological studies on the validity of self-estimates going back nearly 100 years (Shen, 1915). Most of these studies have looked at self-estimates of ability/intelligence (Ackerman & Wolman, 2007; Chan & Martinussen, 2015;

Holling & Preckel, 2005; Gold & Kuhn, 2017; Kang & Furnham, 2016; Tierney & Herman, 1973; Zell & Krizan, 2014) but others have looked at self-estimates of personality (Furnham, 1997, 2001; Ziegler, Danay, Scholmerich, & Buhner, 2010) as well as concepts like emotional intelligence (Siegling, Sfier, & Smyth, 2014). There also have been various reviews (Freund & Kasten, 2012; Furnham, 2016). Some studies have also looked at the ability of people to estimate others, rather than own scores (Cogan, Conklin, & Hollingworth, 1915) while some studies the stability of judgments over time (Jonsson & Allwood, 2003).

The concern of this research is which psychometric test scores are people more and less accurate at estimating and why. This becomes important if self-estimates are used as proxy for actual test scores, under particular circumstances.

This paper reports seven similar studies run on different cohorts that look at the correlations between self- and other-estimated scores and test performance on various preference and ability tests. The methodology allows three issues to be addressed: self-versus other-estimates, self-estimates versus psychometric scores, other-estimates versus psychometric scores.

Self-estimated and psychometrically measured IQ

A limited number of studies have investigated this issue using a fairly diverse series of measures, yet the results have been fairly consistent. De Nisi and Shaw (1977) tested 114 students on 10 different ability tests and also obtained self-ratings. All the correlations between the two scores were significant and positive with five being (r > .30). They concluded that self-reports of ability cannot substitute for validated measures (i.e., IQ tests). Mabe and West (1982) later found a correlation of r = .29 between self-estimated and objective abilities. Borkenau and Liebler (1993) showed that when strangers rated the intelligence of people they saw relatively briefly on a video, the correlation between other-estimate and psychometric score was r = .43.

Various different studies done in Great Britain revealed modest test-estimate correlations (Furnham & Rawles, 1999; Reilly & Mulhern, 1995). Furnham & Chamorro-Premuzic (2004) found a correlation of r = .30 (n = 184) between self-estimates and scores on the Wonderlic Personnel Test. In a cross-cultural study comparing 172 British and Singaporean students, Furnham and Fong (2000) found the correlation between estimated and measured IQ was r = .19 overall (British r = .14; Singaporeans r = .26). The highest correlation was for Singaporean females (r = .51) and lowest for British females (r = .08). In a meta review Zel and Krizan (2014) found a overall r = .29 between self-evaluated and overall performance measures.

Paulus, Lysy and Yik (1998) reviewed the relevant literature and found that correlations between single-item, self-reports of intelligence and IQ scores tended to rarely exceed r = .30 in college students. The authors concluded "as a whole, our verdict is pessimistic about the utility of self-report as proxy measures of IQ in college samples" (p. 551). Recent studies have found similar correlations between self-estimated and psychometrically measured intelligence, thus

supporting Paulus et al.'s (1998) conclusion (Chamorro-Premuzic et al., 2004).

Ackerman and Wolman (2007) thoroughly tested 142 mature American students on a large number of ability tests including verbal, spatial and mathematical tests. Self-estimates were obtained prior to, and after, actual testing. All correlations were positive, though there was wide variability (.27 to .54). Higher correlations were found when both variables were aggregated to make them more reliable: r = 0.33 for spatial ability, and r = 0.44 for mathematical ability. Interestingly participants gave lower estimates for verbal than maths or spatial ability because they had better knowledge of them.

Correlations are affected by two things: whether estimates are made before or after taking the test and which tests are taken. Correlations tend to be more modest (and often more accurate) when self-estimates are made after tests. They also tend to be more modest on crystalised rather than fluid intelligence tests (Furnham, Chamorro-Premuzic, & Moutafi, 2005). In this study we examine participants' ability to predict their own score on a variety of IQ tests to examine the extent to which they vary. It is predicted that both self and other estimated and actual IQ scores would be significantly positively correlated.

Various studies have also looked at emotional intelligence (Petrides & Furnham, 2000; Petrides, Furnham, & Martin, 2004; Siegling et al., 2014). They all found significant positive correlation between self-estimates and actual scores.

Self-estimated and psychometrically assessed personality

There is a small, but consistent, literature on the relationship between estimates of, and scores on, psychometrically validated personality tests. Various studies have looked at participants' ability to predict their own Extraversion and Neuroticism scores (Vingoe, 1966; Harrison & McLaughlin, 1969; Gray, 1972; Semin, Rosch, & Chassein, 1981; Blaz, 1983). Studies in this area have used a large number of personality measures, including the Fundamental Interpersonal Relations Oriented-Behaviour (FIRO-B), the Myers-Briggs Type Indicator (MBTI) (Furnham, 1990), and locus of control measures (Furnham & Henderson, 1983). Furnham (1997) used the NEO-FFI (Costa & McCrae, 1988) to measure the Big Five personality traits, and found participants were best at predicting Conscientiousness (r = .57), followed by Extraversion (r = .52) and Neuroticism (r = .51). They were least good at predicting their Openness-to-Experience score (r = .33) and Agreeableness (r = .39). Furnham and Chamorro-Premuzic (2004) looked at self-estimate and actual test derived scores on all 30 facets of the NEO-PI-R. The most consistent were for the six Conscientiousness scales (range r = .18 to r = .54; mean r = .41). Overall the correlation for six facets (N1, N2, N3, E3, E4, C5) were r > 0.50 while four (N5, O3, A2, A6) were non-significant. They also showed that less Agreeable, Neurotic participants gave lower estimates of their overall intelligence.

Approach to learning

The literature on approaches to learning antedates the research on learning styles and approaches to learning. Whereas the "style" literature is about how different people choose to process material, the "approaches" literature is clearly

much more concerned with motivation and assessment. The issue is how people approach their learning task. Murray-Harvey (1994) notes that both styles and approaches researchers are concerned with the learning strategy that students use which are considered important attributes that they bring to any learning situation.

Most researchers observed that if students were given a text to read that they knew would be examined on some tried to understand, contextualize and comprehend the "big picture" content while others focused on remembering what they thought were the "facts" that they would be examined on. These two very different approaches have been called deep vs. surface approaches. To adopt the deep approach means to achieve a critical understanding and retention of concepts that are integrated into a knowledge schema and used for problem solving. The surface approach is based on a pragmatic short-term memorization of salient facts for examination or repetition.

This study will use the Biggs (1987) measure, which assesses the surface, deep, and achievement-oriented approach to learning. Because this study tested students it was predicted that correlations between self as well as other estimated actual scores would all be positive and significant.

Creativity

A few have investigated the relationship between self- and objectively measured creativity and have used different measures of creativity (Kaufman, 2006; Karwowski, 2011). In one study the correlation between self-estimated creativity and a test score was r = .27 (N = 64) (Furnham, Zhang, & Chamorro-Premuzic, 2006). Because most people believe they are creative and because the concept is so loosely defined we predicted a low, but positive and significant correlation between self-estimates and test score.

This study

A central question is which psychometric test scores people are able to predict with any degree of accuracy. It could be assumed that people are able to predict scores for dimensions that they understand or where they have some frame or schema of reference. If, for instance, a person is required to estimate his or her Extraversion or Conscientiousness score accurately, he or she would have to be familiar with the psychological concept, be clear about the situations or phenomena to which it applied and be aware of how he or she compared with population norms for Conscientiousness and Extraversion. Thus, to do this task well, a participant needs to access and use a cognitive category or framework concerning personality traits.

This study moves the literature forward in three ways. First, while it replicates earlier studies on measures of intelligence and personality, it uses new measures including emotional intelligence, creativity, happiness and approaches to learning. Second, this paper reports seven cohorts of students to examine the replicability of the results. Rather than combine the samples (measured over different years) on those measures, which were the same, we treated this as seven replication studies. Third, we used in all five different measures of intelligence to see if

there were significant differences in the correlations as a function of the different tests.

2. Method

2.1. Participants

Participants were undergraduate students in London. Study 1: N = 72, 55 females, median age 19 yrs; Study 2: N = 95, 71 females, median age 20 yrs; Study 3: N = 91, 74 females, median age 19 yrs; Study 4: N = 118, 90 females, median age 20 yrs; Study 5: N = 106, 85 females, median age 19 yrs; Study 6: N = 102, 71 females, median age 19 ys; Study 7: N = 96, 62 females, median age 20 yrs. All the participants were fluent English speakers and collaborated in this study as part of their course-work.

2.2. Measures

Personality. The NEO Personality Inventory–Revised (NEO-FFI; Costa & McCrae, 1992). This 60-item, non-timed questionnaire which measures the "Big Five" personality factor. The manual shows impressive indices of reliability and validity. Test-retest reliabilities range from r = .71 for Agreeableness to r = .80 for neuroticism.

Approaches to Learning. Study Process Questionnaire (Biggs, 1987). This is a 42-item questionnaire that yields six scores. There are 3 approaches and 2 components. The first component is learning motive (why students learn): the second learning strategy (how students learn). The three approaches are surface (a reproduction of what is taught to meet the minimum requirement), deep (a real understanding of what is learned), and achieving (designed specifically to maximise grade). The questionnaire has been repeatedly shown to have satisfactory internal reliability and test-retest reliability (r = .82), content, construct and predictive validities.

Emotional Intelligence (EQ). Trait Emotional Intelligence (TEIQ) (Petrides & Furnham, 2003). Trait EI "refers to a constellation of behavioural dispositions and self-perceptions concerning one's ability to recognize, process and utilize emotion-laden information. It encompasses various dispositions from the personality domain, such as empathy, impulsivity and assertiveness as well as elements of social intelligence and personality intelligence, the latter two in the form of self-perceived abilities". Studies report test-retest reliability of between r = .74 and r = .84.

Verbal Reasoning. The Baddeley Reasoning Test (Baddeley, 1968). This 64-item test can be administered in 3 minutes and measures Gf through logical reasoning. Scores can range from 0 - 64. Each item is presented in the form of a grammatical transformation that has to be answered with "true"/"false", e.g. "A precedes B – AB" (true) "A does not follow B – BA" (false). The test has been employed previously in several studies (e.g. Furnham & McClelland, 2010) to obtain a quick and reliable indicator of people's intellectual ability. It has a

test-retest reliability of r = .80.

General Knowledge. General Knowledge Test (Von Stumm, 2009). This is a 72 item questionnaire that measures knowledge of six areas: literature, general science, medicine, games, fashion and finance. Each area is measured by 10 items, and each correct response is awarded 1 point (in a few cases, there are two correct responses and not one). The internal reliability of the test for the present sample was a = .78.

Creativity. The Barron-Welsh Art Scale (Barron & Welsh, 1952). This scale consists of 86 different black and white pictures arranged and numbered to 8 pictures per page. Participants are instructed to make quick, instinctive, dichotomous judgements about whether they like/dislike each picture. This test requires no language skills, can be used on children and adults, is simple and does not require extensive concentration. The test-retest reliability is r = .81

Happiness. Oxford Happiness Questionnaire (Hills & Argyle, 2002). It measures trait happiness. This is 29 item scale that was devised the "opposite" of the Beck Depression Inventory. It was one of the first measures to be used in the Positive Psychology revolution and there is a short version. The psychometrics are good though there is some question about its dimensional structure.

Fluid intelligence. Advances Progressive Matrices Set II (Raven, 1938). This is a 36 item test, possibly the most famous in psychology. Participants are shown a diagram with 9 pictures of complex shapes with one missing. Participants have to choose between 8 options of figures that logically fit in the missing space. The test has been extensively validated against other measures of fluid and crystallised intelligence.

General Intelligence. The Wonderlic Personnel Test (Wonderlic, 1990). This 50-item test can be administered in 12 minutes and measures general intelligence. Scores can range from 0 to 50. Items include word and number comparisons, disarranged sentences, serial analysis of geometric figures and story problems that require mathematical and logical solutions. The test has impressive norms and correlates very highly (r = .92) with the WAIS-R total IQ score.

Arithmetic. Mental Arithmetic (Lock, 2008). This is a 30 item test requiring a person to make 10 arithmetic calculations (multiply, divide, add, subtract) per item. It is meant to be a mental test, though some people do attempt written calculations. Ten minutes were allowed for the administration.

2.3. Procedure

Participants in each study were tested simultaneously in a large lecture theatre in the presence of five examiners who ensured the tests were appropriately completed. They completed the tests in two settings each lasting around 40 minutes. Two months later in a lab setting the tests were explained: what each factor measured (i.e. the full definition based on the manuals), and shown population norms and means, as well as the means for their group. They were asked to estimate their (and their friends) score on the same scale shown in the results for each test. For example for the Wonderlic they were shown a normal distribution

scores of over 100,000 showing the range (i.e. 50) the mean score and one standard deviation above and below the mean. They were also given reminders of what the tests looked like to refresh their memory. They were asked to nominate a person in the class who they knew best (i.e. "friend") and also to make an estimate for them. They also indicated on a 5-point scale how well they knew this person from "not much" to "extremely well". This task thus involved around 30 estimates. Immediately after they had completed the exercise they got their test scores, which were explained, in detail. They also saw the correlational results shown in this study two weeks after making their estimates.

3. Results

Study 1: Table 1 presents the descriptive statistics and correlations. Twelve of the 14 self-estimate-actual scores were significant, but only 6 of the other-estimate-actual scores. The highest correlations were for Extraversion and the lowest for Emotional Intelligence. The self-other test scores indicated that the pairs were only significantly alike in their Emotional intelligence, Extraversion and Neuroticism scores. On the other hand their self-other estimate scores indicated that they believed they were alike on 10 scales, particularly General Knowledge and Openness.

Study 2: Table 2 shows the results of the correlational analysis. Of the 14 self-estimated/actual score correlations 12 were significant, but two (Openness and Agreeableness) negative. The highest was for verbal reasoning, followed by deep approach to learning and then Extraversion. Six of the 14 other estimated actual scores were significant, but only two of the self/other actual scores. Ten self/other estimated scores were significant and all positive.

Study 3: Table 3 shows all but one of the self-estimate/actual scores was significantly positively correlated with all intelligence test scores r > .50. Six of the 17 other estimate actual scores were significant. In this study seven of the self and other actual test scores were significantly positive indicating a similarity in personality, ability and approach to learning between the participants. As in the other studies self and other estimates were nearly always significantly and positively related.

Study 4: Table 4 shows with one exception (Openness) all the self-estimate/actual scores were significant with five being r > .50 (Neuroticism, Extraversion, Conscientiousness, Verbal Reasoning and General Knowledge. Nine of the 13 other estimate/actual scores were significant with Extraversion being the highest correlation. Seven of the self/actual scores were significant particularly General Knowledge. All but one of the self/other estimates was significant all being r > .30.

Study 5: In all 11 out of the 13 self-estimate/actual scores were significant the highest being for Happiness, Extraversion and Conscientiousness but with Agreeableness showing a negative relationship (See Table 5). Seven other estimate/actual correlations were significant and two of them were significantly

Table 1. Means and correlations between self and other estimates and test scores.

N = 95	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
	Correlations for the Big Five Personality factors					
Neuroticism	.68**	.35**	.23*	.02	36.38	9.16
Extraversion	.71**	.53**	.33**	.26*	41.64	7.35
Openness	−.36**	−.20	−.04	.63**	29.09	6.06
Agreeableness	−.20	.02	−.02	.58**	29.88	7.28
Conscientiousness	.62**	.43**	.13	.19	43.24	7.16

	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
	Correlations for Approaches to Learning					
Surface	.31**	.14	.15	.58**	43.70	7.33
Deep	.41**	.24*	.22*	.03	47.40	7.09
Achieving	.37**	.01	.15	.33**	47.18	8.49

	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
	Correlations for IQ scores					
EQ	.07	.11	.56**	−.07	4.90	0.70
VR (Baddeley)	.46**	.25*	.12	.28**	31.66	12.08
General Knowledge	.26*	.36**	.07	.69**	32.36	19.73
Creativity	.31**	.10	−.07	.37**	30.99	8.67
Fluid Intelligence	.25*	.12	.20	.25*	12.93	4.32
General Intelligence	.42**	.15	.18	.42**	29.53	7.33

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed).

Table 2. Means and correlations between self and other estimates and test scores.

N = 72	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
	Correlations for the Big Five Personality factors					
Neuroticism	.48**	.28*	.07	.23	38.35	8.45
Extraversion	.56**	.44**	.22	.44**	41.21	6.05
Openness	−.33**	−.38**	.04	.42**	31.49	7.08
Agreeableness	−.43**	−.31**	.16	.24*	29.10	6.60
Conscientiousness	.37**	.35**	.22	−.02	43.34	6.47

	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Correlations for Approaches to Learning						
Surface	.36**	.13	.27*	.55**	46.23	7.69
Deep	.54**	.18	−.16	.12	46.75	7.74
Achieving	.17	−.12	.05	.65**	48.09	6.65

	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Correlations for IQ scores						
EQ	.44**	.14	.30*	.15	140.91	19.39
VR (Baddeley)	.61**	.33**	.01	.25*	33.05	12.77
General Knowledge	.40**	.10	.02	.52**	22.05	36.73
Creativity	.28*	.08	.15	.46**	36.59	7.83
Fluid Intelligence	.36**	.10	.12	.26*	10.27	8.23
General Intelligence	.17	.17	.09	.48**	28.78	5.56

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed).

Table 3. Means and correlations between self and other estimates and test scores.

N = 91	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Correlations for the Big Five Personality factors						
Neuroticism	.20*	.00	.26**	.46**	35.45	7.89
Extraversion	.23**	.18	.22*	.40**	41.75	5.11
Openness	−.47**	−.32**	.29**	.49**	41.85	6.36
Agreeableness	−.42**	−.39**	.00	.47**	42.04	5.07
Conscientiousness	.48**	.35**	.29**	.64**	41.35	6.40

	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Correlations for Approaches to Learning						
Surface	.20**	.00	.26**	.46**	43.80	7.11
Deep	.24**	.34**	.14	.50**	45.76	8.86
Achieving	.05	.00	.20*	.60**	45.48	7.44

	Correlations for IQ scores					
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
EQ	.55**	.12	−.10	.71**	147.30	22.16
VR (Baddeley)	.70**	.30**	.16	.58**	35.08	14.68
General Knowledge	.59**	.27**	.31**	.84**	28.97	10.73

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed).

Table 4. Means and correlations between self and other estimates and test scores.

N = 118	Correlations for the Big Five Personality factors					
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Neuroticism	.59**	.39**	.12	.30**	34.70	8.03
Extraversion	.62**	.55**	.21*	.31**	42.25	6.18
Openness	−.04	−.15	.25*	.01	29.16	5.49
Agreeableness	−.26**	−.19*	.26**	.39**	32.59	5.41
Conscientiousness	.62**	.23*	.17	.44**	42.54	6.99

	Correlations for Approaches to Learning					
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Surface	.44**	.08	−.12	.42**	44.90	6.69
Deep	.45**	.32**	.19	.38**	46.64	6.91
Achieving	.26**	.28**	.30**	.44**	46.09	7.84

	Correlations for IQ scores					
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
EQ	.26**	.30**	.19	.33**	149.11	16.66
VR (Baddeley)	.50**	.38**	.04	.52**	31.59	11.29
General Knowledge	.57**	.44**	.56**	.70**	42.85	15.96
Creativity	.43**	02	.19*	.48**	32.47	9.59
General Intelligence	.31**	.13	.19*	.69**	29.84	5.81

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed).

Table 5. Means and correlations between self and other estimates and test scores.

| N = 106 | Correlations for the Big Five Personality factors | | | | | |
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Neuroticism	.41**	.14	.05	.24*	36.31	8.32
Extraversion	.50**	.21*	.31**	.28**	39.97	6.52
Openness	.47**	−.24*	.10	.37**	31.47	6.62
Agreeableness	−.41**	−.30**	−.36**	−.37**	29.03	6.24
Conscientiousness	.51**	.50**	−.05	.33**	42.17	6.44

| | Correlations for Approaches to Learning | | | | | |
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Surface	.24*	.03	.04	.46**	45.15	7.04
Deep	.41**	.32**	.03	.42**	45.32	7.32
Achieving	.02	.16	.09	.70**	45.57	11.85

| | Correlations for IQ scores | | | | | |
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
VR (Baddeley)	.39**	.19	.12	.65**	36.58	13.21
Creativity	.39**	−.16	.51**	.52**	34.06	12.89
General Intelligence	.15	.32**	.29**	.85**	28.34	5.54
Arithmetic	.46**	.35**	.30**	−.15	13.88	5.54
Happiness	.61**	.11	.11	.08	118.93	18.07

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed).

negative (Openness and Agreeableness). The highest correlation was for Conscientiousness, r = .50. Five self-other actual scores were significant and they showed that this group were alike on their Extraversion, Creativity, general intelligence and arithmetic scores but different on their agreeableness ratings. Finally all but two of the self-other estimate scores were significant.

Study 6: Table 6 shows eleven of the 13 self-estimate/actual score correlations were significant though two were negative (Openness and Agreeableness). Seven of the other estimate/actual correlations were significant and one negative (Agreeableness). Only two of the self/other actual scores were significant suggesting that these participants were not very similar to each other, though nine

Table 6. Means and correlations between self and other estimates and test scores.

| N = 102 | Correlations for the Big Five Personality factors | | | | | |
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Neuroticism	.59**	.33**	.09	.06	38.11	7.86
Extraversion	.55**	.20	−.15	.06	40.55	6.14
Openness	−.37**	−.18	.07	.23*	30.02	6.91
Agreeableness	−.31**	−.29**	.05	.12	28.57	6.44
Conscientiousness	.37**	.26*	.15	.29**	42.00	6.50

| | Correlations for Approaches to Learning | | | | | |
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Surface	.30**	.24*	−.16	.36**	45.12	7.13
Deep	.46**	.29**	.08	.26*	46.27	7.18
Achieving	.19	.11	−.04	.66**	46.07	9.33

| | Correlations for IQ scores | | | | | |
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
EQ	.43**	.11	.24*	.25*	146.63	18.84
VR (Baddeley)	.51**	.19	−.01	.23*	34.47	11.46
Creativity	−.01	−.16	−.15	.33**	33.15	9.41
General Intelligence	.33**	.32**	.03	.20	30.21	5.34
Arithmetic	.60**	.35**	.38**	.61**	14.10	5.50

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed).

of the thirteen self/other estimates were significant suggesting they believed they were.

Study 7: Table 7 shows seven of the 10 self-estimate/actual correlations were significant the highest being for intelligence and one being significantly negative (Agreeableness). Only two of the ten other estimate/actual correlations were significant, one being negative. Three of the self/other correlations were significant.

The data for the Big Five personality factors was aggregated and the analysis repeated. Findings are shown in Table 8. The four columns tell the story of findings in this area. Three of the five correlations were highly significant (r > .45) indicating that participants could predict their Extraversion, Neuroticism

Table 7. Means and correlations between self and other estimates and test scores.

N = 96	Correlations for the Big Five Personality factors					
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Neuroticism	.31**	.09	.37**	.27**	33.09	4.45
Extraversion	.44**	.07	.23	−.04	40.45	4.99
Openness	.22	.08	.11	.33**	36.25	3.58
Agreeableness	−.38**	−.28*	.04	.36**	37.65	4.93
Conscientiousness	.28*	.35**	.29**	.64**	41.35	6.40

	Correlations for IQ scores					
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
EQ	.16	−.02	−.08	.06	124.91	9.49
VR (Baddeley)	.62**	.02	−.26**	.09	31.93	11.22
Creativity	.24*	.09	.14	.42**	32.97	10.07
General Intelligence	.18	−.02	.14	.34**	34.81	5.67
Arithmetic	.59**	.41**	−.14	.59**	10.63	3.85

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed).

Table 8. Means and correlations between self and other estimates and test scores.

N = 670	Correlations for the Big Five Personality factors					
	Self Estimate Self Actual	Other Estimate Other Actual	Self Actual Other Actual	Self Estimate Other Estimate	Self Actual Mean	Self Actual SD
Neuroticism	.47***	.23***	.17**	.23***	36.06	7.74
Extraversion	.51***	.31***	.20***	.32***	40.97	6.04
Openness	−.19***	−.20***	.12**	.35***	32.64	6.02
Agreeableness	−.27***	−.25***	.02	.26***	32.69	6.00
Conscientiousness	.46***	.35***	.17**	.36***	42.28	6.62

*Correlation is significant at the 0.05 level (2 tailed); **Correlation is significant at the 0.01 level (2 tailed); ***Correlation is significant at the 0.001 level (2 tailed).

and Conscientiousness scores. However for the other two factors Openness and Agreeableness the correlations were about half the size and negative. The second column showed the same pattern but correlations were lower, especially for the three positive traits. The third column which looked at personality similarities between the pair four correlations were positive, especially Extraversion, but one

not significant (Agreeableness). The final column showed similar correlations 23 > r > 36 indicating that people believed they were similar to their friends.

Self-Estimates and Test Scores

Summarizing the results from multiple studies using the same measure showed that each sample used an intelligence measure. All seven studies used the Baddeley Reasoning Test. Correlations varied from .38 to .70 with a mean of r = .54. Four studies used the General Knowledge test and results varied from .26 to .59 with a mean of r = .26. Three studies measured fluid intelligence with the Ravens test and correlations varied from .18 to .25 with a mean of r = .26. In all six studies used the Wonderlic test to measure General Intelligence and correlations varied from .15 to .58 with a mean of r = .33. Two studies used the arithmetic test with a mean of r = .54. Six studies included a measure of emotional intelligence and correlations ranged from .07 to .55 with a mean of r = .33.

Six studies used the Approach to Learning measure. The range for self-estimated/actual sores for Deep Learning was r = .31 (Range .20 to .44), Surface Learning was r = .42 (Range .24 to .46) and Achievement Learning r = .18 (Range .02 to .37). Five studies included a measure of creativity. Correlations varied from .00 to .43 with a mean of r = .28. One study included happiness and the correlation between estimated and actual scores was r = .60

4. Discussion

This study showed that with few exceptions people could predict their psychometric test scores. This was true for creativity, emotional intelligence, happiness, intelligence and three of the Big Five personality traits. These results are consistent with previous literature (Semin, Rosch, & Chassein, 1981; Blaz, 1983; Harrison & McLaughlin, 1969; Furnham, 1997). Further, the fact that there were no positive significant correlations between self-estimated and psychometric Openness and Agreeableness scores is also in line with initial predictions. It was hypothesized that the relative obscurity and low usage in the definition/label of these factors would hinder participants' capacity to estimate their scores accurately. However, it is the fact that in many studies, and overall, those correlations were negative suggesting that those who were most Open and Agreeable tended to give low scores and vice versa. Whilst the size of the correlations differed in the four studies, Extraversion showed highest correlations which is to be expected given the way it is so commonly discussed.

Second, regarding the relationship between measured and estimated IQ scores, the present results confirmed our prediction that people would be able to estimate their intelligence to a moderate but significant degree. In fact, the correlation reported between self-estimated and psychometric IQ scores (i.e., r = .30), is not only consistent with the previous literature (De Nisi & Shaw, 1977; Borkenau & Liebler, 1993; Furnham & Rawles, 1999, Zell & Krizan, 2014), and with Paulus, Lysy and Yik's (1998) meta-analysis, in which the authors concluded that self-estimated and psychometrically measured intelligence typically correlate by

(r = .30) (see also Furnham, 2001). The correlations for verbal reasoning (Baddeley, 1968; Furnham & McClelland, 2010) varied from r = .46 to r = .70. This may have been easiest to judge because participants knew how many items they completed which were very familiar. In this sense, power test results may be easier to predict than preference test results.

Apart from the first sample the self-actual estimated correlation for EQ were significant and between r = .26 and r = .55 which confirms Petrides and Furnham (2003). All three studies using the creativity measure showed significant correlations between r = .28 and r = .43 which are similar to the results reported by Karwowski (2011) for creative self-efficacy. This is perhaps surprising given the doubts about the validity of the measures used (Barron & Welsh, 1952), indeed for all creativity tests (Batey & Furnham, 2006). The studies also showed that students were quite able to predict their Approaches to Learning score with Deep Learning correlations higher than Surface Learning. This is not surprising, as students seem to understand this concept very well indeed.

The other person estimate-actual correlations showed three things. First, that correlations were lower (and frequently not significant) compared to self-estimate-actual correlations (study 1: 6/14; study 2: 6/14; study 3: 6/11; study 4: 9/13; study 5:7/13; Study 6:7/13; Study 7:3/10). Despite variation between the studies it seems friends could significantly predict each other's Extraversion, Neuroticism, Conscientiousness, Verbal Reasoning and General Knowledge. However in many studies the correlations were negative and occasionally significantly so. This was reasonably good as they had known each other on average for only three to four months. When partial correlations were computed for the rating of how well they knew the other person there were surprisingly few differences supporting the "thin slices of behavior" research which shows how little information is required for people to make accurate judgments of others.

The self-other actual test scores indicated how similar the two friends/acquaintances were. Some samples showed very few correlations (Study 2) while others yielded more (Study 3) and there was not a detectable pattern. This may be because of the length and nature of the participant's friendship with one another but also supports the similarity-attraction literature which suggests that people with similar personalities, values and abilities are attracted to each other. However most of these participants had only known each other for around three months and many other factors, such as where they lived, may have been a more important predictor of who made friends with whom. There was no clear relationship between how long and well the participant's reported knowing one another and their ability to predict their scores.

The final set of correlations (self-other estimates) replicated previous studies. They showed that most were significant (Study 1: 10/14; Study 2: 10/14; Study 3: 11/11; Study 4: 12/13) being r > .60. It is not clear whether this represented a belief on the part of participants that they really were similar to the friends/acquaintances or whether this was an artifact of the rating style or the requirements of the study where they may have seen the others actual estimates.

The question remains as to who are able to make better estimates than others (i.e. are cognitively or emotionally intelligent people better self- and other-estimators); on what characteristics (personality; ability) they are more or less likely to be accurate; and when (i.e. under what test conditions) they make better estimates. We examined some of these questions: for instance, do cognitively and/or emotionally intelligent people (those scoring over one standard deviation on the tests) do better in predicting their own scores. None of the analyses showed an unequivocally clear pattern. Equally we considered whether those individuals with higher self-actual estimate correlations had a different test profile and again the results were not clear.

These studies however can only be considered as studies of personal awareness to the extent there is considerable evidence of the construct validity of the measures. That is, for the scores to be considered "actual" measures of a characteristic, reliable and valid tests need to be used. Otherwise these studies may be described as those of personal validation of psychometric instruments suggesting that personal estimates are the valid scores that tests hope to correlate with.

Studies such as this may inform various literatures such as that on self-awareness which suggests self-awareness is a highly desirable characteristic in adults.

5. Conclusion

As expected, when they understand a psychological concept most people are reasonably able to predict their own personality test score. It seems the more commonly used the psychological concept, like Extraversion, the more people have an understanding of themselves, although they do not necessarily understand the psychological process or mechanism underlying the test. Typically, correlations between self-estimated and test-assessed scores vary from r = .20 to r = .50 depending on the size of the group, the characteristic being assessed, the personality, sex and culture of the participant and the motive for doing the test. Inevitably, people are far less accurate at estimating the test scores of their friends which depends on how well they know them, understand the concept they are asked to consider and whether their friend sees the score. Overall it seems that self-estimates are only a very weak proxy for actual test scores.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Ackerman, P., & Wolman, S. (2007). Determinants and Validity of Self-Estimates of Abilities and Self-Concept Measures. *Journal of Experimental Psychology, 13,* 57-78. https://doi.org/10.1037/1076-898X.13.2.57

Baddeley, A. (1968). A 3 min. Reasoning Test Based on Grammatical Transformation. *Psychonomic Science, 10,* 341-342. https://doi.org/10.3758/BF03331551

Barron, F., & Welsh, G. (1952). Artistic Perception as a Possible Factor in Personality Style. *Journal of Psychology, 33,* 199-203.
https://doi.org/10.1080/00223980.1952.9712830

Batey, M., & Furnham, A. (2006). Creativity, Intelligence and Personality. *Genetic, Social and General Psychology Monographs, 132,* 455-492.
https://doi.org/10.3200/MONO.132.4.355-430

Biggs, J. (1987). *Study Process Questionnaire Manual.* Hawthorne: Australian Council for Educational Research.

Blaz, M. (1983). Perceive Extraversion in a Best Friend. *Perceptual and Motor Skills, 53,* 891-897. https://doi.org/10.2466/pms.1983.56.3.891

Borkenau, P., & Liebler, A. (1993). Convergence of Stranger Ratings of Personality and Intelligence with Self-Ratings, Partner-Rating and Measured Intelligence. *Journal of Personality and Social Psychology, 65,* 546-553.
https://doi.org/10.1037/0022-3514.65.3.546

Chamorro-Premuzic, T., Furnham, A., & Moutafi, J. (2004). The Relationship between Estimated and Psychometric Personality and Intelligence Scores. *Journal of Research in Personality, 38,* 505-513. https://doi.org/10.1016/j.jrp.2003.10.002

Chan, T., & Martinussen, R. (2015). Positive Illusions? The Accuracy of Academic Self-Appraisals in Adolescents with ADHD. *Journal of Pediatric Psychology, 8,* 1-11.

Cogan, L., Conklin, A., & Hollingworth, H. (1915). An Experimental Study of Self-Analysis, Estimates of Associates, and the Results of Tests. *School & Society, 2,* 171-179.

Costa, P., & McCrae, R. (1988). *The NEO-PI/FFI Manual Supplement.* Odessa: Psychological Assessment Resources, Inc.

Costa, P., & McCrae, R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual.* Odessa: Psychological Assessment Resources.

De Nisi, A., & Shaw, J. (1977). Investigation of the Uses of Self-Reports of Ability. *Journal of Applied Psychology, 62,* 641-644. https://doi.org/10.1037/0021-9010.62.5.641

Dixon, P., Humble, S., & Chan, D. (2016). How Children Living in Poor Areas of Dar Es Salaam, Tanzania Perceive Their Own Multiple Intelligences. *Oxford Review of Education, 42,* 230-248. https://doi.org/10.1080/03054985.2016.1159955

Freund, P. H., & Kasten, N. (2012). How Smart Do You Think You Are? A Meta-Analysis on the Validity of Self-Estimates of Cognitive Ability. *Psychological Bulletin, 138,* 296-321. https://doi.org/10.1037/a0026556

Furnham, A. (1990). Can People Accurately Estimate Their Own Personality Test Scores? *European Journal of Personality, 4,* 319-327. https://doi.org/10.1002/per.2410040405

Furnham, A. (1997). Knowing and Faking One's Five-Factor Personality Scores. *Journal of Personality Assessment, 69,* 229-243. https://doi.org/10.1207/s15327752jpa6901_14

Furnham, A. (2001). Self-Estimates of Intelligence. *Personality and Individual Differences, 31,* 1381-1405. https://doi.org/10.1016/S0191-8869(00)00232-4

Furnham, A. (2016). Whether You Think You Can, or You Think You Can't—You're Right. Differences and Consequences of Beliefs about Your Ability. In R. Sternberg, S. Fiske, & D. Foss (Eds.), *Scientists Making a Difference: The Greatest Living Behavioral and Brain Scientists Talk about Their Most Important Contributions* (pp. 297-230). Cambridge: Cambridge University Press.

Furnham, A., & Chamorro-Premuzic, T. (2004). Estimating One's Own Personality and Intelligence Scores. *British Journal of Psychology, 95,* 149-160.
https://doi.org/10.1348/000712604773952395

Furnham, A., & Fong, E. (2000). Self-Estimated and Psychometrically Measured Intelligence. *North American Journal of Psychology, 2,* 191-200.

Furnham, A., & Henderson, M. (1983). The Mote in Thy Brother's Eye, and the Bean in Thine Own: Predicting One's and Others' Personality Test Scores. *British Journal of Psychology, 74,* 381-389. https://doi.org/10.1111/j.2044-8295.1983.tb01870.x

Furnham, A., & McClelland, A. (2010) .Word Frequency and Intelligence Testing. *Personality and Individual Differences, 48,* 544-546.
https://doi.org/10.1016/j.paid.2009.12.001

Furnham, A., & Rawles, R. (1999). Correlations between Self-Estimated and Psychometrically Measured IQ. *Journal of Social Psychology, 139,* 405-410.
https://doi.org/10.1080/00224549909598400

Furnham, A., Chamorro-Premuzic, T., & Moutafi, J. (2005). Personality and Intelligence: Gender, the Big-Five, Self-Estimated and Psychometric Intelligence. *International Journal of Selection and Assessment, 13,* 11-24.
https://doi.org/10.1111/j.0965-075X.2005.00296.x

Furnham, A., Zhang, J., & Chamorro-Premuzic, T. (2006). The Relationship between Psychometric and Self-Estimated Intelligence, Creativity, and Personality and Academic Achievement. *Imagination, Cognition and Personality, 25,* 119-145.
https://doi.org/10.2190/530V-3M9U-7UQ8-FMBG

Gold, B., & Kuhn, J.-T. (2017). A Longitudinal Study on the Stability of Self-Estimated Intelligence and Its Relationship to Personality Traits. *Personality and Individual Differences, 106,* 292-297. https://doi.org/10.1016/j.paid.2016.10.052

Gray, J. (1972). Self-Rating and Eysenck Personality Inventory Estimates of Neuroticism and Extraversion. *Psychological Reports, 30,* 213-214.
https://doi.org/10.2466/pr0.1972.30.1.213

Harrison, N. W., & McLaughlin, R. J. (1969). Self-Rating Validation of the Eysenck Personality Inventory. *British Journal of Social and Clinical Psychology, 8,* 55-58.
https://doi.org/10.1111/j.2044-8260.1969.tb00585.x

Hills, P., & Argyle, M. (2002). The Oxford Happiness Questionnaire: A Compact Scale for the Measurement of Psychological Well-Being. *Personality and Individual Differences, 33,* 1073-1082. https://doi.org/10.1016/S0191-8869(01)00213-6

Holling, H., & Preckel, F. (2005). Self-Estimates of Intelligence-Methodological Approaches and Gender Differences. *Personality and Individual Differences, 38,* 503-517.
https://doi.org/10.1016/j.paid.2004.05.003

Jonsson, A.-C., & Allwood, C. (2003). Stability and Variability in the Realism of Confidence Judgements over Time, Content Domain, and Gender. *Personality and Individual Differences, 34,* 559-574. https://doi.org/10.1016/S0191-8869(02)00028-4

Kang, W., & Furnham, A. (2016). Gender and Personality Differences in the Self-Estimated Intelligence of Koreans. *Psychology, 7,* 1043-1052.
https://doi.org/10.4236/psych.2016.78105

Karwowski, M. (2011). It Doesn't Hurt to Ask….But Sometimes It Hurts to Believe: Polish Students' Creative Self-Efficacy and Its Predictors. *Psychology of Aethetics, Creativity and the Arts, 5,* 154-165. https://doi.org/10.1037/a0021427

Kaufman, J. (2006). Self-Reported Differences in Creativity by Gender and Ethnicity. *Applied Cognitive Psychology, 20,* 1065-1082. https://doi.org/10.1002/acp.1255

Lock, N. (2008). *The 30 Second Challenge*. London: John Blake.

Mabe, P., & West, S. (1982). Validity of Self-Evaluation of Ability. *Journal of Applied Psychology, 67,* 280-296. https://doi.org/10.1037/0021-9010.67.3.280

Paulus, D., Lysy, D., & Yik, M. (1998). Self-Report Measures of Intelligence: Are They Useful as Proxy IQ Tests? *Journal of Personality, 66,* 523-555.

Petrides, K. V., & Furnham, A. (2000). Gender Differences in Measured and Self-Estimated Trait Emotional Intelligence. *Sex Roles, 42,* 449-461. https://doi.org/10.1023/A:1007006523133

Petrides, K. V., & Furnham, A. (2003). Trait Emotional Intelligence. *European Journal of Personality, 17,* 39-57. https://doi.org/10.1002/per.466

Petrides, K. V., Furnham, A., & Martin, G. N. (2004). Estimates of Emotional and Psychometric Intelligence: Evidence for Gender-Based Stereotypes. *Journal of Social Psychology, 144,* 149-162. https://doi.org/10.3200/SOCP.144.2.149-162

Raven, J. (1938). *Progressive Matrices.* London: Lewis.

Reilly, J., & Mulhern, G. (1995). Gender Differences in Self-Estimated IQ. *Personality and Individual Differences, 18,* 189-192. https://doi.org/10.1016/0191-8869(94)00148-L

Semin, G., Rosch, E., & Chassein, J. (1981). A Comparison of the Common-Sense and Scientific Conceptions of Extraversion-Introversion. *European Journal of Social Psychology, 11,* 77-86. https://doi.org/10.1002/ejsp.2420110105

Shen, E. (1915). The Validity of Self-Estimate. *Journal of Educational Psychology, 16,* 104-107. https://doi.org/10.1037/h0073590

Siegling, A., Sfeir, M., & Smyth, H. (2014). Measured and Self-Estimated Trait Emotional Intelligence in a UK Sample of Managers. *Personality and Individual Differences, 65,* 59-64. https://doi.org/10.1016/j.paid.2014.01.027

Tierney, R., & Herman, A. (1973). Self-Estimate Ability in Adolescence. *Journal of Counselling Psychology, 20,* 298-302. https://doi.org/10.1037/h0034828

Vingoe, F. (1966). Validity of the Eysenck Extraversion Scale as Determined by Self-Ratings in Normals. *British Journal of Social and Clinical Psychology, 5,* 89-91. https://doi.org/10.1111/j.2044-8260.1966.tb00959.x

Von Stumm, S. (2009). *A New General Knowledge Test.* London: UCL.

Wonderlic, E. (1990). *Wonderlic Personnel Test.* Libertyville: WPTI.

Zell, E., & Krizan, Z. (2014). Do People Have Insight into Their Abilities? *Perspectives on Psychological Science, 9,* 111-125.

Ziegler, M., Danay, E., Scholmerich, F., & Buhner, M. (2010). Predicting Academic Success with the Big 5 Rated from Different Points of View: Self-Rated, Other Rated and Faked. *European Journal of Personality, 24,* 341-355. https://doi.org/10.1002/per.753