Personality and ability predictors of the “Consequences” Test of divergent thinking in a large non-student sample

Adrian Furnham a,*, John Crump a, Mark Batey b, Tomas Chamorro-Premuzic c

a Psychology Department, University College London, 26 Bedford Way, London WC1, UK
b Manchester Business School, M15 6PB, UK
c Department of Psychology, Goldsmiths, University of London, UK

ABSTRACT

Over 3000 adult managers attending an assessment centre completed a battery of tests including three personality trait inventories (NEO-PIR; MBTI; and HDS), two ability tests (GMA, WG) and a well established measure of divergent thinking (the Consequences Test) used as the criterion variable for creativity. Regressions showed the NEO-PIR Big Five at facet and domain level accounted for around ten percent of the variance in divergent thinking. The MBTI, Big Four, accounted for only five percent of the total variance. Both intelligence tests were modestly correlated with creativity. Together sex, intelligence and personality accounted for 12% of the variance. Bright, stable, open, extraverted males scored most highly on the measure of creative thinking. Implications of these findings are discussed.

1. Introduction

Most psychometric investigations of the creativity construct have utilised tests of divergent thinking (DT) both historically (see Barron & Harrington, 1961 for a review) and recently (see Batey & Furnham, 2006 for a review). DT tests “require individuals to produce several responses to a specific prompt” (Plucker & Renzulli, 1999, p. 38) in contrast to measures of convergent thinking, such as traditional ability tests, which are based on problems which require the identification of single correct responses. Although both ability and personality correlates of DT have been investigated, past studies have relied predominantly on small student samples and assessed only a limited set of ability or personality traits (notably the Big Five) (Chamorro-Premuzic, 2007; Chamorro-Premuzic & Furnham, 2005). Overcoming these limitations, the current study sought to examine the relationship of a widely used test of DT, namely the Consequences test (Christensen, Merrifield, & Guilford, 1953) with a battery of personality measures and ability tests, using a large non-student sample of working adults. The goal was to understand the nature of the relationship between personality and DT.

1.1. DT and the Consequences Test

Following on from the early work on fluency Guilford (1950, 1967) was one of the first to operationalise creativity in terms of tests of DT, which have been demonstrated to have good predictive validity (Plucker, 1999). This study used the Consequences Test (Christensen & Guilford, 1958; Christensen et al., 1953). It contains a number of questions such as “What would be the consequences if everyone suddenly lost the ability to read and write?” and “What would be the consequences if none of us needed food any more to live?”. Participants are given a specific time either per problem, or for all problems. Responses as for other DT tests may be assessed quantitatively or qualitatively. This is usually done by consensual rating techniques where a pool of expert and/or trained judges make a range of specific judgments with respect to issues like overall quality, originality and realism as well as complexity, use of principles, or the number of positive vs negative outcomes. Perhaps the best known scoring technique is that of Hennessey and Amabile (1988) who specified six principles while others (i.e. Mumford, Marks, Connelly, Zaccaro, & Johnson, 1998) have added further attributes. Table 1 shows a good example.
attraction of the test is that it can easily be adapted for use with school children (Rusch, Denny, & Ives, 1965) but also working adults (Mumford et al., 1998). A major appeal of the Consequences Test is that it does not draw upon existing knowledge (or crystallised intelligence) and therefore it may be used in many different populations. The Consequences Test has been tested with regard to its divergent and convergent validity as well as its concurrent and predictive validity (Gelade, 1995; Vincent, Decker, & Mumford, 2002). The test has been used in various vocational settings to look at such issues as leadership skills and performance (Connelly et al., 2000).

1.2. DT and intelligence

Numerous researchers have examined the relationship between DT and intelligence. Batey and Furnham’s (2006) review of the area concluded that DT and intelligence show correlations in the area of \( r = .20 \) to \( r = .40 \) which suggests a moderate to large effect. This is the case for samples as diverse as architects and air force officers, to ordinary and gifted school children. Effects larger than \( r = .3 \) are considered in the top third of all effects published in psychology (Hemphill, 2003).

Studies have examined cognitive ability (i.e. intelligence) correlates of the DT Consequences Test specifically. Mumford et al. (1998) tested over 1800 military personnel and found measures of intelligence correlated on the order of \( r = .21 \) to \( r = .29 \) with Consequences Test ratings of quality, originality, realism and complexity. In a more sophisticated study of 110 military leaders Vincent et al. (2002) found intelligence correlated on the order of \( r = .25 \). Hence it was predicted that in this study both measures of cognitive ability would correlate with Consequence Test scores (H3).

1.3. DT and personality

The study of the personality traits associated with DT has also been well documented and is reviewed in Batey and Furnham (2006). Research utilising the Eysenck Personality Questionnaire found both positive significant relationships between DT and Psychoticism (Aguilar-Alonso, 1996; Woody & Claridge, 1977) in addition to non-significant relationships (Kline & Cooper, 1986). Significant positive relationships were also observed with regards to Extraversion (Aguilar-Alonso, 1996). Research employing the five-factor model paradigm has tended to find consistent positive relationships of DT to Openness to Experience (King, Walker, & Broyles, 1996; McCrae, 1987; Wuthrich & Bates, 2001) and Extraversion (King et al., 1996). These results received wider support in the comprehensive meta-analysis of Feist (1998). Recently Chamorro-Premuzic and Reichenbacher (2008) found Openness and Extraversion to be the only positive significant Big Five trait predictors of DT but that Neuroticism was a negative predictor.

On the basis of previous research it was predicted that Extraversion (H4) and Openness (H5) would be correlated with DT based creativity. In addition to the Big Five this study also utilised the “Big Four” dimensions derived from the Myers-Briggs. There have been various studies examining the relationship between the MBTI and other measures of creativity (Carne & Kirton, 1982; Jacobson, 1993). It has been argued that creative individuals tend to be more Intuitive rather than Sensing (N > S); Perceiving rather than Judging (P > J); Extraverted rather than Introverted (E > I) and thinking rather than Feeling (T > F) (Thorne & Gough, 1991). On the basis of previous studies it was hypothesised that there would be a positive significant relationship between participant scores on Intuition (H4), Perceiving (H5) and Extraversion (H6).

There is also a recent interest in a two factor higher order categorisation of the Big Five with two factors (Stability and Plasticity). Silvia (2008) suggested that Plasticity is more strongly related to creativity than Stability. This study will attempt to replicate this finding.

There is a growing interest in the relationship between DT scores and “dark-side” measures of personality or sub-clinical measures of psychopathology. Researchers have found relationships between hypomania and DT (Furnham, Batey, Anand, & Manfield, 2008; Schuldberg, 2000–2001) in addition to schizotypy and DT (Batey & Furnham, 2009; Green & Williams, 1999). This is the first study to examine the relationship of the Hogan Development Survey (HDS: Hogan & Hogan, 1997) to DT, based on the HDS manual (Hogan & Hogan, 1997) and other studies on the personality disorders (Oldham & Morris, 1991). It was predicted that various dimensions would be significantly associated with DT: two positive Imaginative (H4), Colourful (H5) and two negative Cautious (H6) and Diligent (H10).

2. Method

2.1. Participants

There were a total of 2603 participants of whom 87% were males. They were all middle to senior managers of various multinational communication organizations. They ranged from their late 30 s to their middle 50 s (mean age = 40.3 yrs; SD = 11.41 yrs).

2.2. Measures

2.2.1. NEO personality inventory form 5 (NEO-PI; Costa & McCrae, 1985)

The NEO personality inventory is based upon the five-factor model of trait personality (Costa & McCrae, 1985). It measures
Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Each single factor/domain consists of six primary factors/facets which can be summed to form a total domain score. The inventory is composed of 240 self-descriptive statements. The manual provides impressive evidence of both reliability and validity. It is probably the most investigated and well used personality test in research based studies that attempt to measure traits (Furnham, 2008).

2.2.2. Myers–Briggs type indicator-form G (MBTI; Briggs & Myers, 1987)

The Myers–Briggs indicator is a Jungian based inventory that uses a paper-and-pencil self-report format. It is composed of 94 forced-choice items that constitute the four bipolar discontinuous scales which are implied in Jung’s theory, Introversion–Extraversion, Sensation–Intuition, Thinking–Feeling and Judging–Perceiving. Respondents can be classified into one of 16 personality types based on the largest score obtained for each bipolar scale (e.g. a person scoring higher on Introversion than Extraversion, Intuition than Sensation, Feeling than Thinking and Judging than Perceiving would be classified as an Introverted Intuitive Feeling Judging). In this study scores on each of the eight scales were used. The Myers–Briggs indicator has evidence that it has satisfactory validity and reliability (Furnham, 2008).

2.2.3. HDS (Hogan & Hogan, 1997)

The HDS was explicitly based on the DSM Axis II personality disorder descriptions, but it was not developed for the assessment of all the DSM disorders. The HDS focuses only on the core construct of each disorder from a dimensional perspective (Hogan & Hogan, 2001, p. 41). An overview of the item selection guidelines can be found in Hogan and Hogan (2001). The survey includes 154 items, scored for 11 scales, each grouping 14 items. Respondents are requested to ‘agree’ or ‘disagree’ with the items. The measure also has a social desirability scale.

2.2.4. The Watson–Glaser critical thinking appraisal (WGCTA; Watson & Glaser, 1980)

This is a timed (40 min) ability test assessing the ability to define a problem, to select pertinent information for its solution, to recognize stated or unstated assumptions, to formulate and select hypotheses, and to draw valid conclusions. The test consists of five subtests: Studies on the WGCTA have provided evidence for the test’s reliability and validity (Watson & Glaser, 1980).

2.2.5. Graduate and managerial assessment; abstract (GMA:A; Blinkhorn, 1985)

This is a timed (30 min) high level test of abstract reasoning ability, which measures the ability to think conceptually, to discover underlying patterns within a set of information, and to switch easily between contexts and level of analysis. The test is made up of 115 questions split into 23 groups of five questions. There are two different scoring methods, the Lenient score (GMA-L), which measures the total number of individual questions that are correct, and the Harsh score (GMA-H), in which a mark is assigned for each group of five questions that are answered correctly.

2.2.6. Consequences (Christensen et al., 1953)

This was the measure of creativity and is discussed in detail in Table 1. Participants were given five items (e.g. What would be the results if people no longer needed or wanted sleep?) They were given 30 min to answer. Repponses were scored by expert judges who were trained and whose inter-rater reliability was frequently tested. Eight consultants met for training days, where they independently scored Consequence Test items, which were first inter-correlated to check reliability and then discussed. This always exceeds 0.80. They came up with two scores total number and total creative scores. Scores represent the total number of non-overlapping original ideas.

2.3. Procedure

Participants from over three dozen different companies in different sectors were required to attend a middle management assessment or development centre where they completed the questionnaires, tests and took part in various exercises. Around 20 trained and experienced consultants scored the tests following standard procedures. The assessment was aimed at determining the suitability of each manager for promotion. Each manager was given feedback on the results, including how he/she related to the test norms as well as his/her colleagues.

3. Results

Because data collection was done at an assessment centre it is possible that scores may be distorted through impression management processes. This could lead to defensiveness on the part of participants with truncated scores and reduced variance. Examination of both sets of scores suggests this may have occurred but only for a few dimensions of the preference tests (i.e. the measure of Neuroticism) but that all other scores were normally distributed around population norms.

3.1. Consequences and the big 5

Table 2 shows the correlational results of the big 5 at the domain and facet level. The correlations show four of the Big Five were correlated with the consequences score. Open, Extraverted,
Stable people scored highest. The regression results also shown in this table revealed that the Big Five accounted for 10% of the variance. This confirms H² and H³.

Table 2 also shows the correlational and regression results for the six facets of each of the three most important factors as shown from the results in Table 2. For the correlations nearly all of the facets were significant with five of the six Neuroticism facets being negative. The regression (with the 18 facets as predictor variables and DT as criteria) showed that they also accounted for 10% of the variance. All six Openness facets were significant predictors (particularly Fantasy) though one (Aesthetics) was negative. Thus Openness contributed the most variance to predict DT however at the domain level it was clear that Extraversion and Openness contributed almost equally to DT.

Two factors were then computed from the five domain scores: Stability (Reversed N & A & C) and Plasticity (E & O). These scores were then correlated with DT (r = .26, N = 3308, p < .001) but Stability was not (r = .02, ns).

3.2. Consequences and the MBTI

Table 3 shows that three of the Big Four MBTI factors were related to creativity particularly the SN dimension. The Regression showed that Perceiving, Intuitive, Extraverts did best and that this accounted for 5% of the variance.

3.3. Consequences and the HDS

Table 4 shows the results of the correlational and regression results for the HDS. Nine of the eleven scales were significantly correlated with creativity particularly Colourful and Imaginative which are both makers of Openness which is an established correlate of creative thinking.

3.4. Consequences and intelligence

Table 5 shows the correlations with intelligence. Both intelligence tests correlated r = .12 with performance on the Consequences Test.

3.5. Regressions

A series of step-wise regressions were then performed using three blocks; Sex, Intelligence and Personality. Variables were entered in different sequences to explore the incremental variance. Table 6 shows the results of the regression where first sex, then intelligence was added 12.3% of the variance was accounted for. This change in R-square was significant (p < .01). The three most powerful significant predictor variables were Extraversion, Openness and Intelligence measured by the Watson–Glaser Test. When IQ was added after personality in a different regression personality accounted for 11% of the variance and intelligence 1.5%.

A similar regression to that of Table 6 was done this time using the MBTI four scores instead of the NEO five. The final regression was significant (F(7, 1829) = 10.75 but only accounted for 4% of the variance. The E/I dimension (B = .05, β = .06, t = 2.38) and J/P (B = .05, β = -.06, t = 2.58) indicated Perceiving Extraverts scored highest. When the “dark side” variables were entered in a step wise regression, after sex and intelligence, the regression, was significant (F(17, 388) = 3.41), and accounted for 11% of the variance. Those high on imagination (B = .03, β = .13, t = 2.33) and low on diligence (B = -.13, t = 2.31) did best.

Finally, a step-wise multiple regression was computed with all the significant correlates of DT as predictors and DT as criterion. Ability and domain personality factors were both included, which decreased the actual sample size to N = 360 because of missing or incomplete data and that less than 500 people completed the dark side measure. In all, the regression accounted for 12% of the variance was accounted for. Table 6 shows that when personality was added 12.3% of the variance was accounted for. This change in R-square was significant (p < .01). The three most powerful significant predictor variables were Extraversion, Openness and Intelligence measured by the Watson–Glaser Test. When IQ was added after personality in a different regression personality accounted for 11% of the variance and intelligence 1.5%.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>β</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>.37</td>
<td>.09</td>
<td>3.68***</td>
</tr>
<tr>
<td>GMA</td>
<td>-.20</td>
<td>.06</td>
<td>2.32*</td>
</tr>
<tr>
<td>WG</td>
<td>.15</td>
<td>.10</td>
<td>4.12***</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>-.05</td>
<td>-.09</td>
<td>3.51***</td>
</tr>
<tr>
<td>Extraversion</td>
<td>.11</td>
<td>.18</td>
<td>7.01***</td>
</tr>
<tr>
<td>Openness</td>
<td>.08</td>
<td>.14</td>
<td>5.46***</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>-.03</td>
<td>-.04</td>
<td>1.79</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.01</td>
<td>-.02</td>
<td>0.44</td>
</tr>
</tbody>
</table>

F(8,1868) = 33.98, R² = .12. ** p < .05. *** p < .001.
variance and the significant predictors retained were: Extraversion \((\beta = .23, t = 4.51, p < .01)\); GMA \((\beta = .19, t = 3.74, p < .01)\); Diligent (HPD) \((\beta = -.15, t = 2.82, p < .01)\); and Imaginative (HPD) \((\beta = .12, t = 2.39, p < .05)\). All other predictors were non-significant when these four factors were taken into account. Structural equation modeling was attempted and results can be obtained from the fourth author.

4. Discussion

All of the hypotheses set out in this study were confirmed by the correlational results. The significant correlations rarely exceeded \(r = .20\) suggesting that these individual difference factors have a modest relationship with DT as measured by the Consequences Test \((\text{Hemphill, 2003})\).

The regresional results competing the NEO-PIR, MBTI and HDS suggested that it was the former that accounted for most of the variance. Further the results were in accord with all previous research in the area \((\text{Chamorro-Premeuzic & Reichenbacher, 2008})\), which indicated three of the Big Five domains related to DT. Open, stable, extraverts scored highest. However the facet analysis showed most clearly that it was Openness facets, particularly the fantasy facet that was the best predictor of DT though Extraversion facets were equally important. This confirms the work in this area \((\text{McCrae, 1987})\).

The final step-wise regressions, even if carried out on a much smaller sub-sample for which complete scores on all measures was available, showed that there is a great degree of overlap between different factors from different inventories when it comes to explaining individual differences in DT. Hence the desirability of Structural Equation Modeling in future studies. Indeed, of all the significant correlates of the Consequences Test only four were sufficient to explain 12% of the variance). The factors that were retained as significant predictors came from the NEO (Extraversion) and the HDI (Diligence and Imaginative), and also added cognitive ability (GMA). The results are congruent with a larger body of previous evidence, derived from separate investigations, suggesting that higher DT is partly a reflection of higher Extraversion, higher Imagination/Openness and lower Diligence (Conscientiousness), as well as higher cognitive ability. In this sense we know the profile of creative thinkers.

Whilst this study had the advantage of having a very large \(N\) and individual scores on many different ability and non ability measures, it had the limitation of having only one measure of creativity. Whilst there is good reason to believe that participants saw this as an ability (power) it is known that different measures of creativity intercorrelate relatively weakly \((\text{Batey & Furnham, 2006})\). The test conditions (an assessment centre) may have advantaged extraverts and thus over emphasised this factor which is not thought of as a very important variable in creativity. Ideally it would have been desirable to have a wider battery of creativity measures as well as more tests of divergent thinking itself. It would also be desirable to test participants under less “competitive” conditions such as an assessment centre.

References