

Using Visual Questionnaires to Measure Personality Traits

—VisualDNA—

June 23, 2014

Abstract

Personality is a strong predictor of real-life outcomes and behaviours. Personality assessment has been employed in numerous fields of academic as well as commercial settings, and enables the prediction of a wide range of behaviours in both the physical world as well as the online space. VisualDNA has developed an innovative online image-based assessment methodology to measure personality. Visual questionnaires offer a fun, interactive and gamified experience, while maintaining the validity of the outcome measures. The methodology was applied to create a portfolio of reliable personality assessment tools. The tests have been applied in multiple commercial settings and have successfully predicted valuable outcomes for businesses. This paper describes the theoretical foundations and validation process of our assessments

1 Introduction

A vast body of psychological research relates real life outcomes to variance in personality traits. Examples range from academic achievement (Chamorro-Premuzic Furnham, 2003), job performance (Barrick Mount, 1991, 2001), subjective well-being (Weiss, Bates, and Luciano (2008), health and longevity (Kern Friedman, 2011) and a range of other real-life outcomes (Ozer Benet-Martinez, 2006; Roberts et al., 2007).

VisualDNA is a leading commercial provider of psychological insights at scale. Hundreds of thousands of users completed our propriety visual psychological questionnaire. Hundreds of millions of users are then profiled using a behavioural inference algorithm. In this paper we focus on the first part and demonstrate the rigorous methodology we use to construct fun,

interactive and gamified experience, while maintaining the validity of the outcome measures.

1.1 The “Big Five” dimensions of personality

Psychologists have studied personality extensively over the past several decades and many theories have been developed regarding its definition and the traits that it encompasses. A major consensus has been reached in the 1990s (Digman, 1990) with the Five Factor Model (FFM) exemplified by Costa McCrae’s work (1987). Norman (1963) and Tupes Christal (1961) have been regarded as the original fathers of the FFM which states that individual differences in personality can be categorised into five major traits: Neuroticism, Extraversion, Agreeableness, Conscientiousness and Openness to Experience. These traits, also known as the “Big Five”, have become universally known and are used by researchers and practitioners alike: there is in fact substantive evidence for the use of the FFM as a framework to describe individual differences in personality which are generalisable across cultures (Costa McCrae, 1992; McCrae John, 1992). Although different taxonomies and labels exist for each of the five personality dimensions, the model is generally agreed upon and is used as the main classification of personality traits. These traits are defined as “personal dispositions that are stable over time and that influence a person’s patterns of behaviours in different situations” (Chamorro-Premuzic, 2007). The following behaviours are typically associated with the five traits:

- *Openness to experience.* Imaginative, cultured, curious, original, broad-minded, intelligent, and artistically sensitive.
- *Conscientiousness.* Being careful, thorough, responsible, organized, and planful. In addition, hardworking, achievement-oriented, and persevering.
- *Extraversion.* Sociable, gregarious, assertive, talkative, and active.
- *Agreeableness.* Courteous, flexible, trusting, good-natured, cooperative, forgiving, soft-hearted, and tolerant.
- *Neuroticism.* Anxious, depressed, angry, embarrassed, emotional, worried, and insecure.

One of the established scales that measure the “Big Five” personality dimensions is the International Personality Item Pool (IPIP) — Five Factor Model (3). The IPIP items are presented as brief and clear behavioral

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate
1. Am the life of the party.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Feel little concern for others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Am always prepared.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Get stressed out easily.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Have a rich vocabulary.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Don't talk a lot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Am interested in people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Leave my belongings around.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Am relaxed most of the time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Have difficulty understanding abstract ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1: Example of IPIP questionnaire ¹.

statements preceded by active verbs. The self-applicability of each statement is rated on a 5-points Likert scale which records the level of agreement or disagreement with a series of statements.

Figure 1 provides an extract from the IPIP test. It illustrates some of the challenges from a UX perspective: it is lengthy, repetitive and very “dry”. VisualDNA has developed an alternative way to measure the “Big Five” traits which offers a fun, interactive and gamified experience to the user. This is accomplished by replacing the textual based responses with vivid images - compare for example Figure 1 with Figure 2 (an example of a VisualDNA question). This enables us to obtain highly positive user engagement; for example, half of our traffic is viral, and completion rates range from 70

The use of image based response options has advantages beyond user experience. Section 2 reviews both the advantages and challenges of constructing an image based questionnaires. Section 3 provides details regarding our methodology. In section 4 we provide the empirical evidence to the accuracy of the “Big Five” questionnaire. We conclude in section ??.

How would you make the most of a morning off? Pick one to start

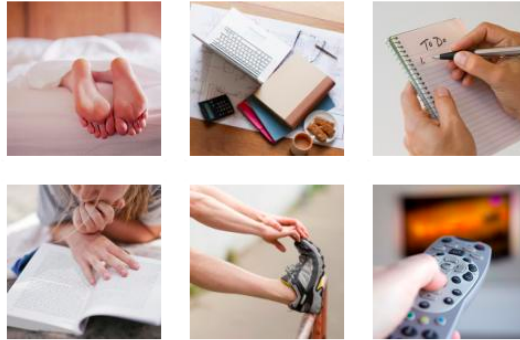


Figure 2: Prototypical question in a VisualDNA's quiz.

2 Using Images in Psychometric Assessment

VisualDNA has developed a visual quiz that uses images (visual response options) in order to measure personality. The potential benefits of using visual questionnaires to measure personality traits in psychometric assessment are summarised below.

- The use of visual tests in psychometric assessment is less subject to 'social desirability bias' because it makes more difficult to the respondent to distinguish between correct and incorrect answers in comparison to traditional Likert scales. With regard to people's reports about themselves, social desirability bias is generally defined as providing responses that are perceived as more acceptable than the response that the respondent would have made under neutral conditions. This response bias occurs mainly for items or questions that deal with personally or socially sensitive content and results from respondents trying to answer questions as a good person rather than in a way that reveals what they actually believe or feel. Social desirability interferes with the interpretation of tendencies as well as it is one of the most common sources of bias affecting the validity of experimental and survey findings.
- The use of images in psychometric testing has the potential to convey more complex ideas, concepts, and emotions, and to move the respon-

dent to a deeper understanding than is possible through words alone. Furthermore, the use of images involves overcoming many barriers to communication that may lead to the message becoming distorted, and conveying a clear and concise message.

- The use of visual tests require significant less effort and attention than traditional psychometric tests, resulting in a reduction of the effect of boredom and fatigue. Traditional psychometric tests are based on very long, repetitive, textual-based questionnaires that create a barrier for respondents who are not psychologically minded. Thus, the completion rate of visual-based tests is potentially much higher than the completion rate of textual-based tests. Overall, this encourages greater engagement with questions.

As can be understood from the above, the potential benefits involving the use of images to measure personality traits are remarkable and opens the door for further opportunities. However, several difficulties and challenges arise in the process of using images in psychometric assessment.

- It is very difficult to maintain uni-dimensionality across response options. In visual-based tests, it is necessary to control the adjustments of the images such as color, contrast or brightness, as well as any extraneous connotations likely to be evoked by images.
- The use of images in psychometric assessment is possibly subject to the same difficulties and confusions in scoring as other 'forced-choice scales'. The use of forced-choice formats has been criticised on the grounds that respondents are provided with only a limited number of response alternatives to choose from (2). Unlike Likert scaling which is a bipolar scaling method, measuring either positive or negative response to a statement, with a forced-choice scale, by contrast, the opposition of meaning between supposedly opposite items cannot be examined. It has to be assumed.

3 Quiz Development Process

Each question in our quiz is designed to facilitate understanding of a distinct personality trait, and is usually composed of a range of possible answers. For example, the question: "How much energy do you have after meeting new people?" was designed to measure Extraversion, with illustrations of batteries merely replacing the standard Likert scale (see Figure3).



Figure 3: Example for a question in which we use a visual metaphor to represent scale.

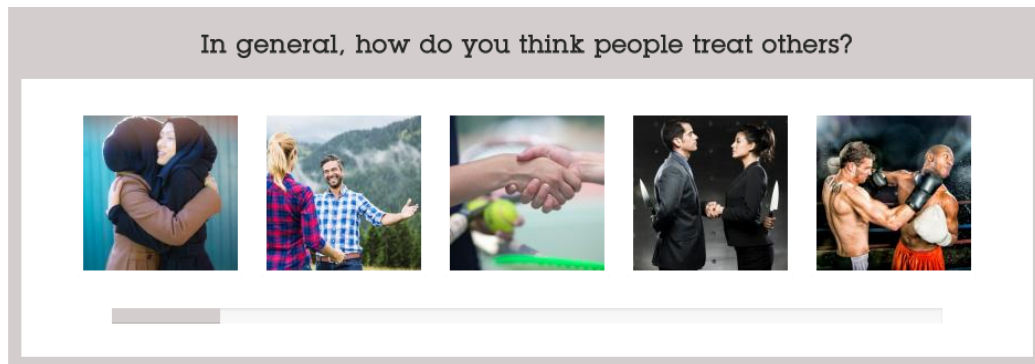


Figure 4: Example for an image based questions.

However, in the next example: “In general, how do you think people treat others?”; the use of images is more sophisticated, and enables us to include variety of positive and negative feelings portrayed through the valence of the images, while enhancing the ease of identification with the chosen answer (see Figure 4).

In practice, we follow an iterative process to come up with the final set of questions. The process typically involves the following steps:

1. The *Qualitative Psychology team* researches various behaviours which are known to be linked with the relevant traits. Together with the *Creative team*, the set of behaviours are transformed into candidate questions and answers. Only questions which are both relevant and suitable for image based representation are selected to the next step.
2. The *Creative team* follows with an image research step in which answers are replaced with corresponding images. Some questions, for which

suitable images are not found are dropped.

3. The final stage involves a panel test, run by our *Quantitative Psychology team*, in which both the clarity of the images and, more importantly, the validity of our scores are tested.

We iterate these steps until the quality of our scores is sufficiently high. In the case of the "Big Five" quiz we ended up with 32 image based questions.

It is of a special interest to explain how we derive and validate our scores. As mentioned in section 2 it is not trivial to generate scores to image based questionnaires using traditional methods. We decided to adopt a computational method instead. In the panel (step 3) we let people answer both our novel visual quiz as well as a standard one. We then use linear regression to assign, algorithmically, appropriate weights for the different images such that the correlation between our score and the standard score is maximised. We provide detailed explanation of this in the next section.

4 Quiz Validation Process

A study was conducted to measure the correlation of our new quiz comprised of 32 image questions, with an established scale of the Big Five personality dimensions taken from the International Personality Item Pool (IPIP, Goldberg, 1992, Goldberg, 1999; Goldberg, Johnson, Eber, Hogan, Ashton, Cloninger Gough, 2006). This method was undertaken to ensure convergent validity of our scale. Correlations of 0.4 - 0.7 imply a strong convergent validity, while correlations above 0.7 signify very strong convergence (Chamorro-Premuzic Ahmetoglu,2013), and are rarely achieved even in test-retest reliability tests of personality scales.

A demographically balanced sample of 1,000 paid participants was recruited for the study through a third-party market research company (see Table 1 for a breakdown). After identifying missing data and filtering out the incomplete/invalid cases, data were analysed from 980 participants. Each of the participants completed VisualDNAs image-based quiz as well as the IPIP one.

Since each respondent completed both tests, we could treat the problem of computing a scale for VisualDNA quiz as a supervised learning problem where the objective is find 'weights' for different images such that the correlation with the IPIP scale is maximised.

Regression models were trained with 10-fold cross validation. The reported correlations (Table 2) were calculated based on predicted scores, and

Table 1: Age and Gender Breakdown of Respondents

Percentage of Respondents	
Male	50.00%
Female	50.00%
18 - 24	16.20%
25 - 34	15.60%
35 - 44	18.00%
45 - 54	16.20%
55 - 64	14.40%
65+	19.60%

therefore provide a conservative measure in comparison to academic literature standards. Figure ?? gives the equivalent plots of the actual IPIP scores and the predicted ones.

Factor	Correlation
Openness	.601
Conscientiousness	.622
Extraversion	.747
Agreeableness	.561
Neuroticism	.711

Table 2: Correlation between VisualDNA and IPIP scale. In all cases the p-value < 0.01 .

5 Conclusions

The Big Five dimensions of personality is an established way of measuring personality traits with a vast amount of studies the demonstrate direct relation to real life outcome. VisualDNA developed an alternative way of measuring the Big Five dimensions of personality which offer a fun, interactive and gamified experience - crucial in commercial settings.

A large-scale validation test was carried out according to academic and industry standard verification protocols and which largely concluded that VisualDNAs image-based assessments are a robust measure of personality.

We use this quiz to provide a unique experience to our users with a robust insight into their personalities. As the barrier of entry to our quizzes is much

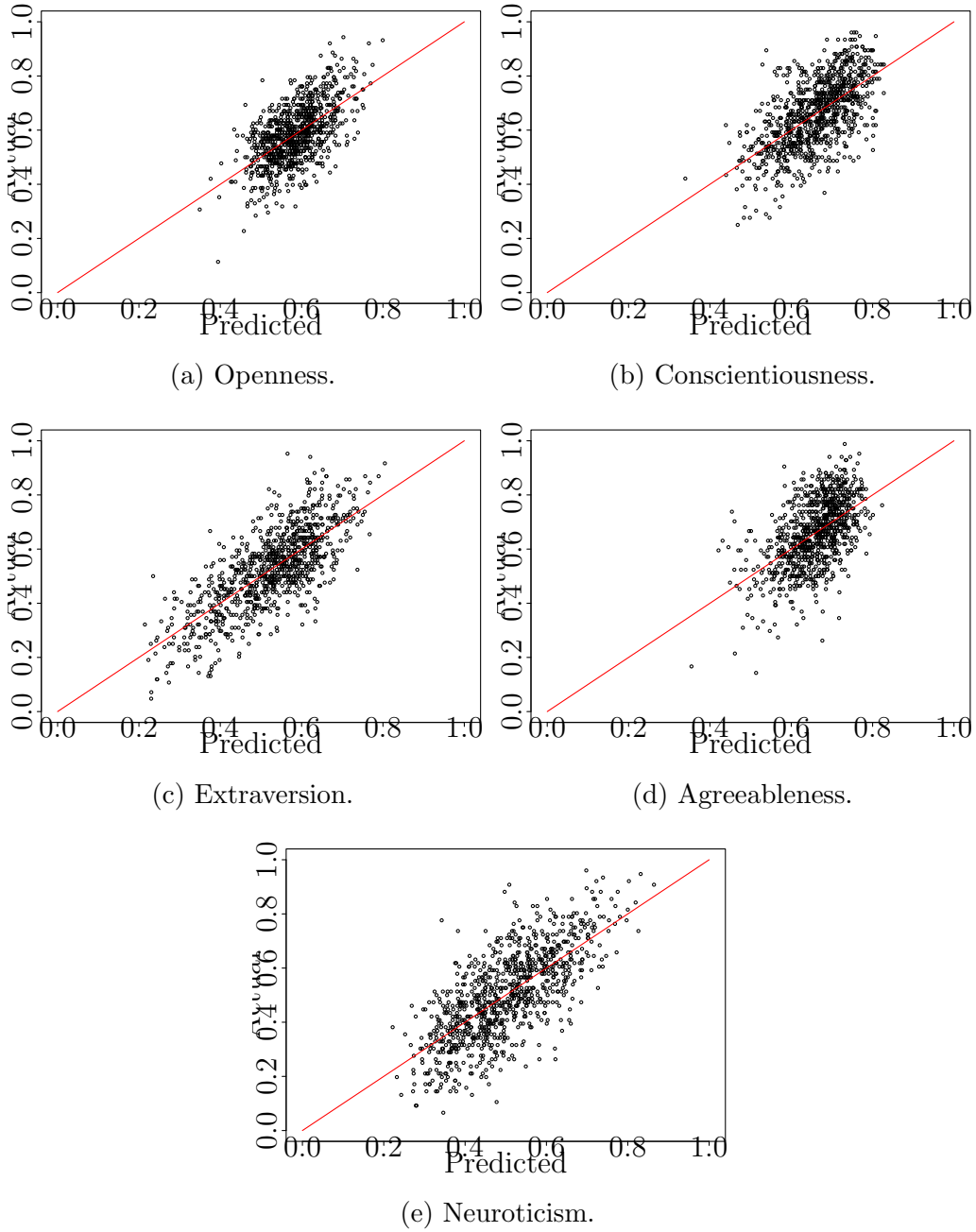


Figure 5: Actual score computed from standard IPIP questionnaire versus predicted score computed from visual quiz for each of the Big Five personality traits. The red line indicates perfect linear correlation.

lower than standard questionnaires we are able to provide these insights to a considerably large number of users.

In a subsequent paper we will demonstrate how we use this quiz as a seed to a much larger study which enables us to infer personality traits to millions of users who did not necessarily completed this quiz.

References

- [1] Lanman, R. W., & Remmers, H. H. (1954). The Preference and Discrimination indices in forced-choice scales. *Educational and Psychological Measurement*, 14(3), 541-551.
- [2] Harrigan, J., Rosenthal, R., & Scherer, K. (Eds.). (2008). *New Handbook of Methods in Nonverbal Behavior Research*. Oxford University Press.
- [3] Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe*, 7, 7-28.