7 Personality traits

7.1 Introduction

The previous chapter dealt with the validity of GMA (general mental ability) as a predictor of job and training performance, as well as other work-related outcomes. In the current chapter, we discuss the predictive power of personality traits.

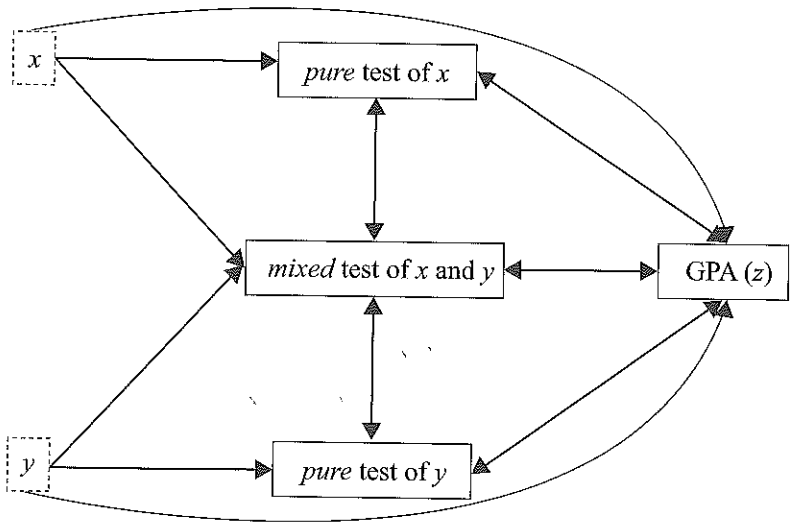
The question then emerges as to what is the difference between GMA and personality traits, and this is a question for which only one simple answer exists: traditionally (in personnel selection as well as in the wider context of psychological assessment), GMA is measured or *tested* via objective performance tests (such as those discussed in Section 6.6 and 6.7), whereas personality traits are *assessed* via subjective inventories, notably self- or other-reports (but especially self-reports). In that sense, one can distinguish between cognitive abilities and personality traits on the basis of assessment methods, whereby the former reflect individual differences in the capacity to identify correct responses to a standardised test (verbal or non-verbal), whereas the latter reflect individual differences in general behavioural tendencies, assessed only subjectively, that is, through people's accounts (one's own or others'). This led to a now well-established distinction in psychology to refer to cognitive abilities in terms of *maximal performance* and personality traits in terms of *typical performance* (Cronbach & Gleser, 1965), though in the case of personality traits 'behaviour' is a more accurate term than 'performance'.

As much as this distinction is straightforward, things get more complex when we try to assess whether measures of ability and inventories of personality may or not be tapping into the same underlying constructs. That is, is the distinction between personality and abilities only relevant at the level of measured constructs, and therefore a purely methodological issue, or are personality inventories and ability tests also assessing similar constructs? Let us chose a simple example to illustrate this question. Let us say that we want to measure individual differences in running, specifically how fast people can run. One option would be to test a number of subjects in a 100 metre race and time how fast they run (using a stopwatch). Another option would be to *ask* people how fast they can run, and we would be willing to bet that at least 90 per cent of our readers would find the first option better than the second. Why? There are two main reasons. First, people may be unaware of how fast they can run. Second, even if they were aware, they

may be unwilling to tell us (especially if we were giving out medals or cash to the fastest runners). Thus, they may chose to exaggerate how fast they can run and try to deliberately mislead us. Of course, one may ask other people (their friends or indeed their enemies) how fast our candidates can run, but the same problems apply, i.e., the friends may be as unwilling or unable to provide this information as the runners themselves, and enemies can hardly be expected to be more accurate.

Our readers may have guessed that these two problems also apply to personality inventories, as personality traits are only assessed subjectively. Indeed, this has been the most common objection to using personality scales in personnel selection and, accordingly, a large part of this chapter is devoted to this issue. But let us make clear at the outset that there are some advantages in using subjective[1] reports rather than objective performance tests. In fact, these advantages highlight some of the limitations of objective assessment methods (briefly discussed in Section 6.13) and explain the importance of using personality inventories in personnel selection. The first issue is that subjective reports can take into account aggregate data, that is, how people have been performing or how they perform most of the time (this has already been shown in Chapter 3 on biodata). Thus, in the context of our running example, asking the runners or people who are familiar with them how fast they can run may provide information on how fast these people *usually* run, as well as how fast they have managed to run in the past. If we were only interested in how fast people *can* run, this may or not be a good approach. However, it is clear that if we were interested in how fast they *tend to run*, then objective performance tests would be very poor indicators of this: even Olympic-medal winners don't run as fast as they can most of the time.

There is also a second issue, which is that objective performance tests (i.e., timing people once, or testing their ability once) may not be accurate, especially if factors other than running ability interfere. Examples of such factors can range from fatigue or test-anxiety (including fear of evaluation, which may explain why even professional athletes may record faster times in training than in the actual competition) to a heavy hangover. This phenomenon, simply referred to as *underperformance* (understandably, 'over-performance' is rarely an issue), threatens the validity of objective tests but does not harm subjective reports. In fact, there are two main reasons underlying the fact that psychometric tests do not perfectly predict any outcomes: the first is that they may be failing to measure the construct in a completely accurate manner; the second is that, even if they do, that construct may be only one of the determinants of the outcome we wish to predict. In a sense, these two reasons are the same and merely one: tests that capture not just running ability but also fatigue, test-anxiety and hangover (to stay with the above example) should therefore predict running performance better than tests that capture only one of the predictors. This logic has been applied to personality inventories and ability measures (Chamorro-Premuzic & Furnham, 2006a; Wechsler, 1939). As shown in Figure 7.1, measures of academic performance, such as general point average (GPA), can be used to validate both



Note: x = 'actual' cognitive ability, y = 'actual' trait anxiety, GPA = grade point average, x and y are predictors, z is the criterion. Since z is influenced by both x and y, tests of x and y will correlate higher with z than *pure* tests of x or y only. ▭ measure ▭ latent ⟷ correlation ⟶ influence

Figure 7.1 *Ability and non-ability determinants of grade-point average (GPA) (from Chamorro-Premuzic & Furnham, 2006a)*

ability and personality scales. Given that GPA is determined by both ability and personality, tests that capture ability as well as non-ability variance should be stronger predictors of either pure ability or pure personality scales alone.

But we were trying to determine whether personality and ability measures may be assessing similar or at least related constructs, rather than deciding which method is more valid. The truth of the matter is that the only way to determine what any psychological instrument is assessing is to use external validity criteria, and, ironically, that is the very problem with psychological assessment, namely the fact that there is no absolute or ultimate criterion. Thus it is always possible to run faster or slower next time, and in fact very unlikely that anyone runs exactly as fast both times. Moreover, because any variable – whether running ability or personality – correlates with and predicts a wide range of other variables/outcomes, the choice of criterion is rarely unquestionable.

With regard to job and training performance, which have been the criteria *par excellence* in personnel selection for over a century, it is interesting that although GMA is a good predictor of job and training performance, we use maximal performance measures (ability tests) to predict typical performance (aggregate levels of achievement at work, for instance income or occupational level). Further, ability tests rarely account for more than 30 per cent of the real variance in job performance (Schmidt & Hunter, 1998), which is one of the

Observation: Watching what people do most of the time (unfeasible and expensive). Interviews and biodata are largely used to 'observe' (infer) people's personalities.

Recommendation letters: Used extensively. Though not explicit personality tests, they provide information on the personality of the candidate/applicant/etc.

Psychometric inventories: Most common way of assessing personality traits, mainly self-reports but can also be completed on behalf of others (other-reports).

Projective tests: Assume that people's unconscious motives or traits are 'projected' (spontaneously expressed); require interpretation from an expert.

Objective tests: Range from physiological measures (transpiration, heart beat, etc.) to response latency (reaction times in self-report responses); long but failed quest.

Situational tests: Putting people in specific situations to see how they react: e.g., in front of others to see if they can lead; in tempting situations to see if they cheat/steal.

Figure 7.2 *Ways of assessing personality traits*

main reasons why in the past fifteen years a growing number of researchers have examined the predictive validity of subjective inventories. The other main reason is adverse impact, as group differences in personality traits are less pronounced than in GMA (see Section 6.5). But the crucial point is that it would make little sense to assess personality if personality inventories were assessing the same constructs as ability tests (Chamorro-Premuzic & Furnham, 2004). Accordingly, even if personality and ability measures are related, they should be independently related to external criteria.

So, what do personality inventories assess, if not intellectual abilities, and, more importantly, why should personality inventories predict job performance if they do not assess cognitive abilities? This chapter will address these two questions.

It should be noted that personality traits are not only assessed via self-report inventories (though that is indeed the most common way of assessing them). Observation, situational tests, projective techniques and even objective measures can also serve as measures of personality, albeit not explicitly (see Figure 7.2). The most commonly used forms of observation in personnel selection are interviews (see Chapter 2) and biodata (see Chapter 4). Employers may not explicitly state that what they are assessing in an interview or looking for in biodata is indeed traces of personality, but there is longstanding evidence for the fact that candidates'/interviewees' personality traits affect employers' decisions (Wagner, 1949), despite the fact that most interviewees fake (Levashina & Campion, 2007). When employers search biodata records they are looking for markers or consequential outcomes of personality traits. In fact, an individual's background or past would mean very little unless it contributed to predicting what he or she will do in the future (as seen in Chapter 4). Hence consistency underlies the motivation

to search for people's biographical records and use that information for organisational decision making. Although biodata measures do overlap with personality inventories (see Section 4.8), they have been shown to have incremental validity over personality (as well as GMA) in the prediction of work performance (Mount, Witt & Barrick, 2000). At the same time, personality inventories have been found to show incremental validity over and above biodata in the prediction of similar outcomes (McManus & Kelly, 1999). Situational tests are a way of 'economising' observations, creating situations that may trigger the behaviours one wishes to observe. Thus they can also be regarded as consequential outcomes of personality traits, though there is longstanding evidence for the fact that personality inventories are only modestly correlated with single behaviours (at a given point in time) (see Box 7.1). On the other hand, the problem with letters of recommendation (and other reports in general) is that they are provided by people who tend to have an interest in the candidate/applicant or at least reasons for being biased in favour of the candidate. Finally, projective tests – which can range from the Rorschach inkblot test (see Section 1.8) to interpretation of dreams – have vanished quite rapidly since the 1970s, although certain techniques, such as graphology (a projective test whereby candidates 'project' aspects of their personality in their writing styles) are still widely employed in certain countries (see Section 1.2), such as France and Argentina, despite evidence that it does not work (Ben-Shakhar, Bar-Hillel, Bilu, Ben-Abba & Flug, 1986; Furnham, Chamorro-Premuzic & Callahan, 2003).

Although it has long been argued that personality traits – or indeed any psychological construct – should not be assessed only with one method, e.g., self-report or interview (Prien, 1977), but with a multi-method approach (Marsh, Martin & Hau, 2006), most researchers and practitioners continue to rely on single methods of assessment and, in the case of personality, that method is self-report inventories. However, it is important to disentangle the variance that is caused by the method of assessment and the actual trait or construct that is being assessed. This is a complex theoretical and methodological issue: for example, self-reports of cognitive ability tend to correlate with performance tests of cognitive ability only at $r = .40$ (Chamorro-Premuzic, Moutafi & Furnham, 2005), meaning they share less than 20 per cent of the variance. This could be indicative of the fact that self-reports and performance-based tests are actually assessing different constructs (and there is wide acceptance that the latter capture the real essence of cognitive ability because they are correlated more highly with performance outcomes). However, with regard to personality traits, objective tests tend to be validated against self-reports, and self-reports are often validated against other-reports (though these are rarely substantially intercorrelated).

Self-report inventories are what people usually call 'personality tests', namely different statements or questions that are answered on a 'yes/no' or 'strongly disagree' to 'strongly agree' response scale. These questions or items are then used to compute overall factors, giving people different scores along different personality dimensions or categorising them in terms of one personality type or another (e.g., 'Extraverted' vs 'Introverted'). What all these variations have

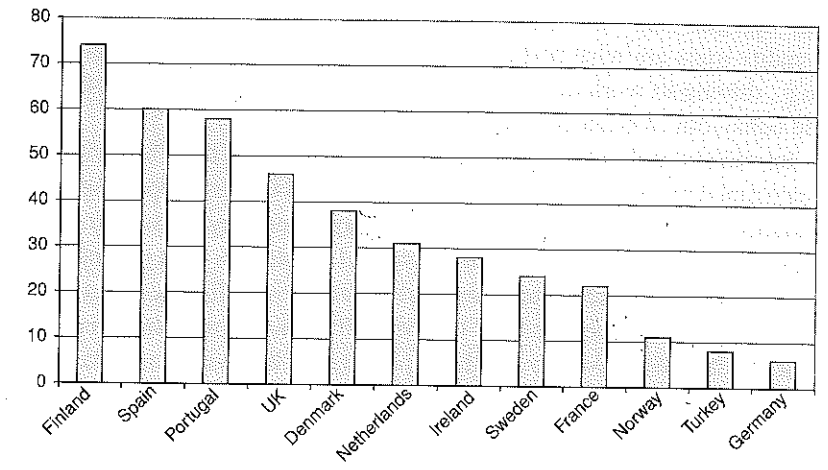


Figure 7.3 *Percentage of companies using psychometric tests in Western Europe (Price Waterhouse Cranfield data, 1994)*

in common is that they assume that respondents are capable of and interested in providing the relevant information about themselves, and that the scores will have cross-temporal reliability (be relatively stable across time) and predictive validity (predict external outcomes, such as job performance). There are four major ways to assess the quality or *construct validity* of personality inventories: (a) *face validity* (in simple terms, how valid do the scales look?); (b) *content validity* (asking questions that genuinely assess the construct we want to assess); (c) *predictive validity* (generating scores that relate to external outcomes); and (d) *factorial validity* (identifying and replicating the same latent dimensions via the same set of items). These four criteria can be applied to the various ways of assessing personality traits specified in Figure 7.2.

The Price Waterhouse Cranfield survey (Dany & Torchy, 1994) found that psychometric testing was used only infrequently in most Western European countries. As shown in Figure 7.3, psychometric tests were used in selection widely only in Finland, and to some extent in Spain and Portugal, but in the remaining Western European countries most companies did not use them and in some countries (Norway, Turkey and Germany) they were barely used at all. US surveys indicate that American organisations (at least in 1997) used personality scales only infrequently. A survey of 251 US employers who were asked to rate how frequently, on a five-point scale (1 = never, 3 = sometimes, 5 = always), they used personality inventories for personnel selection showed the average for both graduate and general jobs to be lower than 2 (closer to 'never' than 'sometimes').

7.2 What are personality traits?

Personality traits are widely defined as stable, inner, personal dispositions that determine relatively consistent patterns of behaviour (including feelings

and thoughts) across different situations (Chamorro-Premuzic, 2007). Thus personality traits are what make a person's behaviour similar in different situations (e.g., driving to work or watching a film with friends), as well as different from the behaviour of other people.

Intuitively, individuals are inclined to believe that their behaviour is largely determined by the situation, and there are clear examples supporting this idea. For instance, few people behave in the same way when they are with friends as when they are with strangers, or, indeed, potential employers (in a job interview). That said, there are marked interindividual differences (differences between people) in how individuals behave both with their friends and with strangers – including in a job interview. Moreover, these patterns of behaviour are likely to emerge every time we are with our friends, every time we are with strangers and every time we have a job interview. Thus, although most people are more at ease in the company of friends than in the company of strangers, and although most people are more anxious during a job interview than while watching a film with friends, *some people* are clearly more comfortable interacting with strangers than others, even in a job interview. Likewise, *some people* are characterised by a tendency to worry all the time, even when they are watching a film with friends (in fact, if they are watching a film just before a job interview they may be completely unable to concentrate on the film), whereas *other people* rarely feel worried or anxious about things (in fact, if they are watching a film before a job interview they may be so relaxed that they forget about the interview).

The key to understanding personality traits is understanding both how people differ and how they are similar. In that sense, personality traits are the cause of behavioural differences among individuals in a given situation, and since these differences are maintained across a number of situations, personality traits should predict how people will behave in comparison to other people most of the time. Accordingly, behaviour can be understood as a product of both situational and personal variables (see Figure 7.4).

Situational variables, which have been traditionally studied within social psychology and under laboratory or experimental conditions, make behavioural differences among people less noticeable. An example of a situational variable affecting behaviour in the workplace may be the prospect of a pay rise (reward) or demotion (punishment), both of which should motivate employees to work more. However, not all employees may be equally responsive to this situation; some may be more likely to work more if they are threatened with demotion, whereas others may only be motivated to increase their productivity if they are tempted with a pay rise. Accordingly, the behaviour we are trying to predict, namely *work output* or productivity, is determined by both the situation (the prospect of rewarding or punishing employees) and personality traits (whatever makes one person more responsive to the prospects of reward and punishment than another person). Thus behaviour (b) is determined (=) by situational (S) and personality (P) factors. Moreover, because the effects of one situation will be different for one person than another (see Figure 7.4), the correct formula to

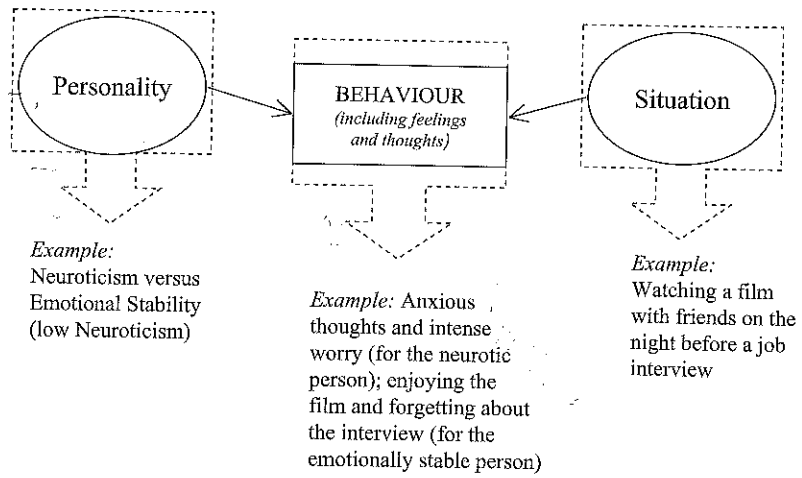


Figure 7.4 *Behaviour as a function of both personality and the situation*

represent the relationship between b, S and P is that of an interaction, namely $b = S \times P$. The multiplication sign illustrates the fact that, if either S or P has a value of 0, the other variable will have no effect on b. For example, an employee may have no ambition of getting a pay rise and may not care about demotion either, which would mean that neither the prospect of a pay rise or a demotion would change his or her work output. Conversely, if an employer is only doing the job for the money and is worried about not earning enough, he or she will be very responsive to the situation.

Box 7.1 Situationalism: undermining personality traits

During the 1960s, the notion of personality traits came under heavy scrutiny when psychologists (especially Walter Mischel) pointed out that personality traits are very weak predictors of behavioural outcomes because behaviour is essentially inconsistent (Mischel, 1968). Thus Mischel concluded that measures of personality (e.g., Extraversion) rarely correlate with behavioural criteria (e.g., being talkative) by more than $r = .30$. This conclusion had a strong negative impact on the research community and led to stagnation in personality research that was also noticeable in the field of applied psychology and personnel selection. However, Mischel and proponents of *situationalism*, i.e., the idea that behaviour is explained better in terms of contextual factors and weakly in terms of inner dispositions, overlooked the importance of personality traits as predictors of aggregate or long-term behavioural outcomes (Eysenck, 1980). Just as it is hard for weather forecasters to make accurate predictions about the weather until one or two days in advance, personality inventories provide poor information of their behavioural consequences at a specific given time. Like the weather,

behaviour does change from one moment to the next. Yet weather forecasts are usually accurate at predicting that the temperature in *x* month will be much warmer than in *y* month, and that, during *x* month, the south will be warmer than the north (Hogan, 1996). Likewise, personality inventories are useful for making predictions of broad behavioural outcomes, especially when situational factors are taken into account (see Section 5.1). In short, personality traits are general constructs and their criteria should be equally general. When this 'bandwith fidelity' is taken into account, trait inventories have shown more than adequate reliability throughout the lifespan, with fifty-year interval test–retest correlations of .25, and usually higher for shorter time periods (Hogan, 1996). Moreover, a meta-analysis reported average test–retest reliabilities for personality in the region of .75 (Viswesvaran, 2000).

7.3 The Big Five: emergence of the five-factor model

The central issue in personality research is not what personality traits are or whether personality traits exist or are useful, but, rather, *which* personality traits should be assessed. In the past twenty years, psychologists have provided compelling evidence for the fact that individual differences in personality can be classified on the basis of five major traits, namely Neuroticism/Emotional Stability, Extraversion/Introversion, Agreeableness, Conscientiousness and Openness to Experience (sometimes referred to as Intellect, Autonomy, or simply 'Factor V'). Indeed, many psychologist believe that these traits capture the essence of interindividual variability by providing a general level of description of the person (P rather than S), which, if valid, should help us predict how people will behave in the future. Although it is questionable whether the Big Five offer the best classification of individual differences (Block, 1995), they are the common currency and universal language in personality research and can be converted or translated into many different taxonomies. Figure 7.5 shows how several different taxonomies and labels can be organised neatly along the five major personality factors. As can be seen, different systems (each row represents a different taxonomy[2]) can be organised quite coherently on the basis of the five personality dimensions lined up at the top (Extraversion, Agreeableness, Conscientiousness, Openness and Neuroticism). Traditionally, this description has not included individual differences in cognitive ability, though in the past ten years a number of researchers have argued that an integration of ability and non-ability traits is needed to provide a more comprehensive picture of the determinants of academic and occupational performance and related outcomes (Ackerman & Heggestad, 1997; Armstrong, Day, McVay & Rounds, 2008; Chamorro-Premuzic & Furnham, 2005; Matthews, 1999).

It is noteworthy that although personality traits were neither identified to predict behaviour at work nor for the purpose of personnel selection, and even though

Extraversion	Agreeableness	Conscientiousness	Openness	Neuroticism
Dominance Initiative	Social orientation	Task orientation		Emotional orientation
Low ego control		High ego control	Ego resiliency	
Activity–Sociability		Impulsivity	Emotionality	
Exvia vs Invia	Pathemia vs Cortertia	Super Ego Strength	Independence Subduedness	Adjusted vs Anxiety
Positive Emotionality		Constraint	Absorption	Negative Emotionality
Extraversion	Psychoticism			Neuroticism
Outgoing, Social, Leadership	Self-protected Orientation	Work Orientation	Aesthetic Intellectual	Dependence
Ambition and Sociability	Likeability	Prudence	Intellect	Adjustment
Extraversion vs Introversion	Feeling vs Perceiving	Judging vs Perceiving	Intuition vs Sensing	Agreeableness

Figure 7.5 *Big Five as universal language of personality. Different rows indicate different researchers' systems*

personality is supposedly unrelated to cognitive ability, the fact that personality traits describe general and consistent differences between individuals means they are useful also in the context of personnel selection, as predictors of work-related behaviours. Moreover, and as mentioned in Section 7.1, if personality traits assess individual differences that are largely independent from cognitive abilities, any information they provide may actually help us improve on the prediction of work-related behaviours beyond cognitive abilities. This is hardly surprising given that people's performance at work is not a 'one off' event but a long-term succession of events where typical behaviour matters more than maximal performance.

Box 7.2 No aversive impact

Hogan (1996) argued that personality inventories generally do not systematically discriminate against any ethnic or national group, and thus may offer more equitable bases for selection (see also John, 1991). There are some sex differences in major personality traits, but these are unlikely to have powerful effects on real-world outcomes and likely to cancel each other out: for instance, women are more neurotic but also more conscientious than men (Feingold, 1994). In addition, there are no personality differences between disabled and non-disabled people, and at least in some personality inventories older individuals (over 40) tend to obtain more favourable or adjusted profiles, probably because they are indeed more mature and balanced. This has no doubt contributed to the

growing use of personality inventories in selection. However, evidence suggests that adding personality inventories alongside GMA measures in personnel selection does not seem to compensate for the adverse impact of GMA tests (Ryan, Ployhart & Friedel, 1998).

7.4 Validity of personality traits as predictors of job and training performance

The question of whether personality inventories should be used or not in the context of personnel selection has divided practitioners and researchers for decades. Practitioners tend to assign much more weight to personality than to abilities, but are reluctant to accept the validity of self-reports because common sense indicates that people can and will fake. On the other hand, researchers are still debating whether faking is really a problem and whether the validities of personality inventories are acceptable, meaningless or high. Moreover, unlike with GMA (where higher scores are always better), the same personality traits may be advantageous for some situations but disadvantageous for others, which means that the effects of personality on work-related outcomes are a lot more context-specific and moderated by situational factors than in the case of GMA.

Thus the answer to the question of whether personality tests should be used in personnel selection will depend mostly on who you ask, even if answers are based on exactly the same data. Even if we agree on the fact that personality matters in the workplace, one has to identify exactly what aspects of personality matter in every job, whether we can assess them accurately, and how significant an effect they have on work and training performance; all of these issues have been debated extensively. What is beyond debate is that personality inventories are weaker predictors of job and training performance than are cognitive ability tests (Ghiselli, 1966; Schmidt & Hunter, 1998). In fact, arguments against personality inventories are rarely based on social policy (which, as seen in Section 6.8, is the basis of the common objection to using cognitive ability tests) and mostly based on methodology, namely how personality traits are assessed. Another common criticism of personality traits is based on the magnitude of their associations with job and training performance outcomes. But how valid are personality traits?

Prior to the consolidation of the five-factor model (especially after the first reviews were published in the 1950s and 1960s), researchers and practitioners were generally inclined to believe that personality inventories had very weak validities in personnel selection. Early meta-analytic evidence in support of the validity of personality traits in predicting job performance suggested that, if the right traits are assessed in the right context, personality inventories are significantly, albeit modestly, correlated with work-related outcomes. For instance, one of the first quantitative reviews of different inventories across different jobs

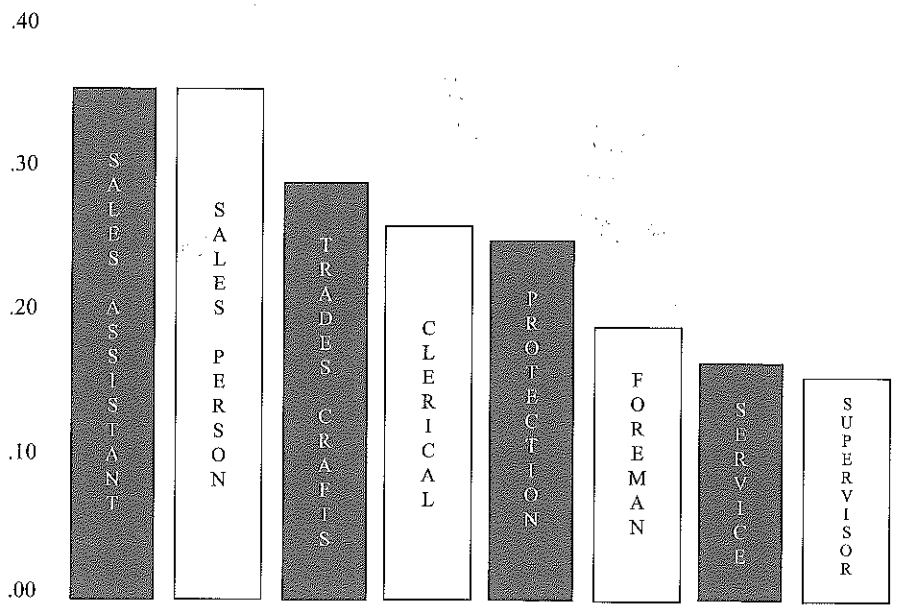


Figure 7.6 *Validity of personality traits across occupations (early meta-analytic evidence) (based on Ghiselli & Berthol, 1953)*

reported uncorrected correlations ranging from .14 to .36 (Ghiselli & Barthol, 1953). As shown in Figure 7.6, the lowest validities were found for supervisory jobs whereas the highest validities were found for sales jobs. However, subsequent reviews during the 1950s and 1960s found that in most studies (80–90 per cent) personality traits did not significantly predict work-related criteria. Until the 1990s, however, there had been little consensus as to the structure of personality, which meant that almost every other researcher had his or her own 'language' for describing major dispositions. This, plus the general lack of quantitative reviews (most reviews were descriptive rather than meta-analytical), contributed to the little support for the idea that self-report inventories assessing stable and general dispositions could be predictive of important work outcomes.

The emergence of the Big Five model provided occupational and I/O researchers with a universal language to compare the results from different validity studies. To be clear, personality experts have and will continue to disagree as to which traits best describe individual differences (Block, 2001), but by the 1990s sufficient evidence had accumulated for the existence of five distinct factors across a wide range of cultures, languages and instruments (Digman, 1990). The first meta-analytic review of findings on the validity of the 'Big Five' (the authors did not refer to these traits as such and also looked at three additional constructs) was provided by Hough and colleagues (Hough, Eaton, Dunnette, Kamp & McCloy, 1990). Results are summarised in Figure 7.7. As shown, personality scales were quite consistently related to individual differences in physical fitness and military bearing (the coefficients shown did not correct for restriction of range,

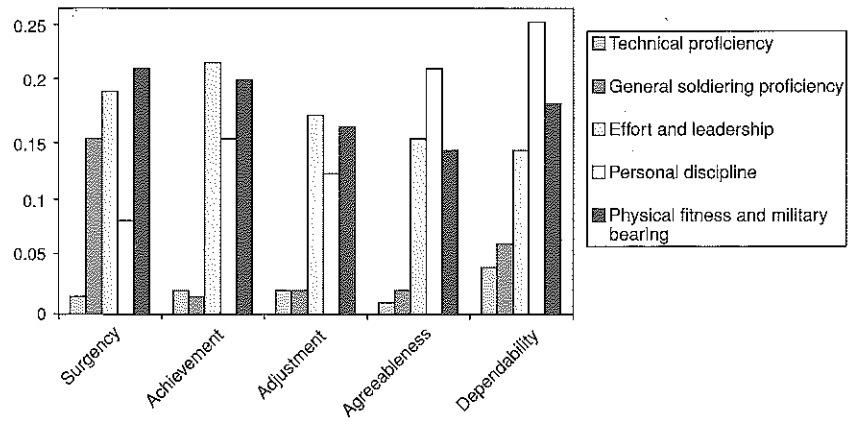


Figure 7.7 *Validities for ABLE personality traits (based on Hough* et al., *1990)*

unreliability in the predictor or criterion). There were also several significant associations between personality traits and indicators of effort and leadership, and personal discipline, but the validity of traits in predicting technical proficiency and general soldiering proficiency was considerably lower. Looking at individual traits (Surgency represents Extraversion, Achievement and Dependability represent Conscientiousness, and Adjustment represents Emotional Stability/low Neuroticism), we can see that all traits, especially Extraversion and Conscientiousness, were related to occupational outcomes. Although the authors did not estimate the 'true validities' for these traits (at the level of actual constructs), it is safe to assume that the values would have been much higher.

Although the results of Hough *et al.* added support to the idea that personality traits are valid predictors of job performance, the authors reported data from a single instrument (the ABLE inventory) and military occupations only. Two years later, however, Barrick and Mount's (1991) now widely cited quantitative review meta-analysed the results from 117 investigations (162 samples and 23,994 participants) published between 1952 and 1988, organising their findings according to the five personality traits. The authors looked at five job families, namely professionals, police, managers, sales and skilled/semi-skilled jobs. Results provided compelling evidence for the predictive power of Conscientiousness scales across different settings (predicting job and training performance) and the validity of other traits in specific contexts (for instance, Extraversion was a predictor of managerial and sales jobs, and Openness was a good predictor of training success).

Figure 7.8 depicts Barrick and Mount's validities for the Big Five across different occupational criteria (combining all different job types). As shown, productivity, turnover/tenure, status change, salary and especially subjective ratings where best predicted by Conscientiousness. Indeed, other Big Five traits were inconsistently linked to these five criteria. For example, Openness was related to status change and turnover/tenure but not to salary, subjective ratings and

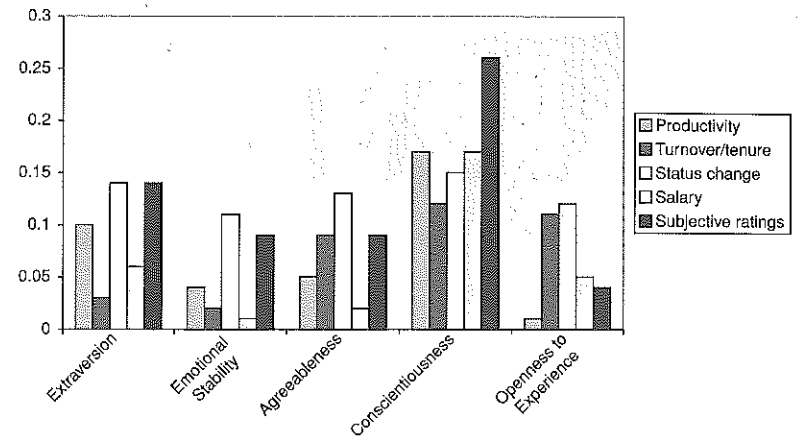


Figure 7.8 *Meta-analysis of Big Five predicting objective and subjective work criteria (based on Barrick & Mount, 1991)*

productivity, whilst Extraversion was linked to status change and subjective ratings but not to turnover/tenure or salary. It is also fair to say that the validities were all modest, even in the case of Conscientiousness, which was by far the best personality predictor of job performance. This led I/O psychologists (both researchers and practitioners) to continue, perhaps even increase, the debate on whether personality inventories should be used in personnel selection, though Barrick and Mount's own conclusion was that the use of personality scales in personnel selection was warranted.

The same year as Barrick and Mount's meta-analysis, Tett and colleagues (Tett, Jackson & Rothstein, 1991) meta-analysed data from 494 studies (97 samples and 13,521 participants), including not only the Big Five but also Type A personality (which combines individual differences in status-consciousness, impatience, restlessness and pro-activity) (Friedman, Hall & Harris, 1985) and locus of control (the tendency to attribute events to personality or situational factors) (Ryckman & Rodda, 1971). Their results replicated several of the findings from Barrick and Mount. Furthermore, Tett *et al.* found that confirmatory studies yielded even higher validities for personality traits as predictors of job performance than thought. However, there were also some 'befuddling' inconsistencies (Goldberg, 1993) between these two meta-analyses. For example, Tett *et al.* reported a validity of .27 for Openness and job proficiency, whereas Barrick and Mount's was −.03. Moreover, the latter found that Conscientiousness was the most solid and powerful predictor of job outcomes but the former found that three other traits (Emotional Stability, Openness and especially Agreeableness) worked better. There were also some discrepancies between these two studies and other meta-analyses. However, Mount *et al.* (2000) and Tett *et al.* justified these discrepancies in terms of differences in statistical methods of analysis and goals between the studies (Ones, Mount, Barrick & Hunter, 1994; Tett, Jackson, Rothstein & Reddon, 1994).

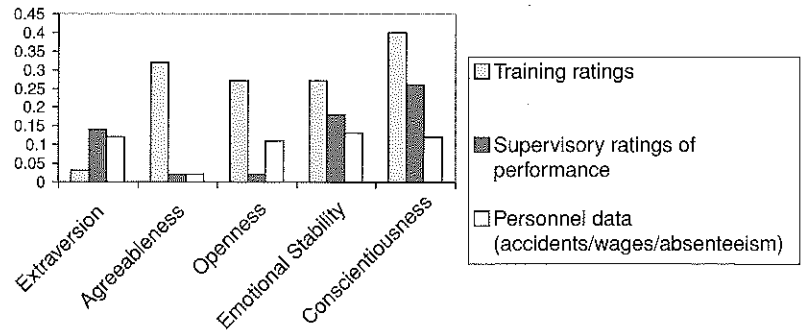


Figure 7.9 *Personality and job performance in the EC (validities from Salgado's meta-analysis)*

7.5 Personality in the European Community

Mount *et al.*'s and Tett *et al.*'s meta-analyses were based on US data. A few years later, a meta-analysis of European validity studies of personality traits by Salgado showed that not only Conscientiousness, but also Emotional Stability (low Neuroticism) had generalisable validities predicting job outcomes (Salgado, 1997). Uncorrected correlations for personality ranged from .00 to .15. The highest observed validities were found for Conscientiousness (.15), Agreeableness (.12), Emotional Stability and training (.11), and Openness (.11) as predictors of training success. Although these coefficients are very modest, the 'true validities' (correlations at the construct level) were at least twice as high (e.g., .39 for Conscientiousness and training, .31 for Agreeableness and training, and .27 for Emotional Stability and training). Moreover, true validities for several traits and supervisory ratings approached these coefficients. Figure 7.9 illustrates the comparative validity of the Big Five (according to Salgado's EC meta-analysis) as predictors of training ratings, supervisory ratings of performance, and 'personnel data' (e.g., absenteeism, salary and accidents at work). As seen, Conscientiousness showed the highest predicted validities across the three criteria – with moderate predictive validity. The figure also highlights the fact that Extraversion was a very poor predictor of training ratings, that Agreeableness was unrelated to both supervisory ratings of performance and personnel data, and that Openness was unrelated to supervisory ratings. Moreover – with exception of Extraversion – the Big Five had higher validities for training than for other criteria.

Whether the results from the US and European meta-analyses should encourage the use of personality inventories in personnel selection or not has been a matter of extensive debate. What is clear, however, is that there has been an exponential increase in research into the validity of personality traits in I/O psychology since these meta-analytic studies and the five-factor model appeared.

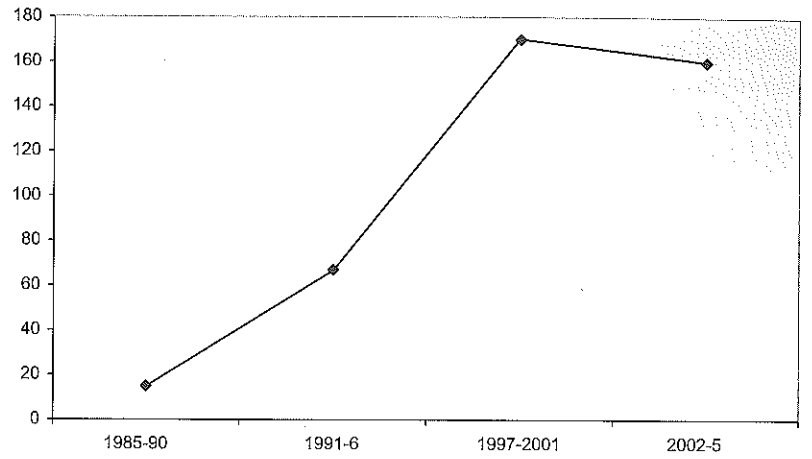


Figure 7.10 *Publications related to personality and selection between 1985 and 2005 (based on Morgeson* et al., *2007)*

Figure 7.10 (based on Morgeson, Campion, Dipboye *et al.*, 2007) illustrates the increase in the number of validity studies on personality reported in the main journals of applied psychology and the Society of Industrial and Organisational Psychology. As seen, there were fewer than twenty studies until 1991 (when Barrick and Mount's meta-analysis was published), and three times that amount between 1991 and 1996, and another near-threefold increase between 1997 (the date of Salgado's meta-analysis) and 2001, with a similar number of studies between then and 2005 (the absolute number has dropped slightly but the relative number has gone up as there are two less years in this period).

7.6 Conscientiousness: the most important personality predictor of work outcomes

Regardless of where one stands in relation to the use of personality inventories (criticisms are discussed in Section 7.18), it is clear that Conscientiousness is the most important personality predictor of job performance (Schmidt & Hunter, 1998), and thus the most important non-ability factor in personnel selection, at least among the Big Five personality traits. It has therefore been argued that Conscientiousness is equivalent to GMA in the realm of personality and self-reports. This is hardly surprising given that Conscientiousness assesses individual differences in responsibility, dutifulness, achievement-striving, organised planning and self-control (see Figure 7.11). Thus conscientious people are more likely to both set themselves ambitious goals and work hard to accomplish them (Barrick, Mount & Strauss, 1993; Gellatly, 1996).

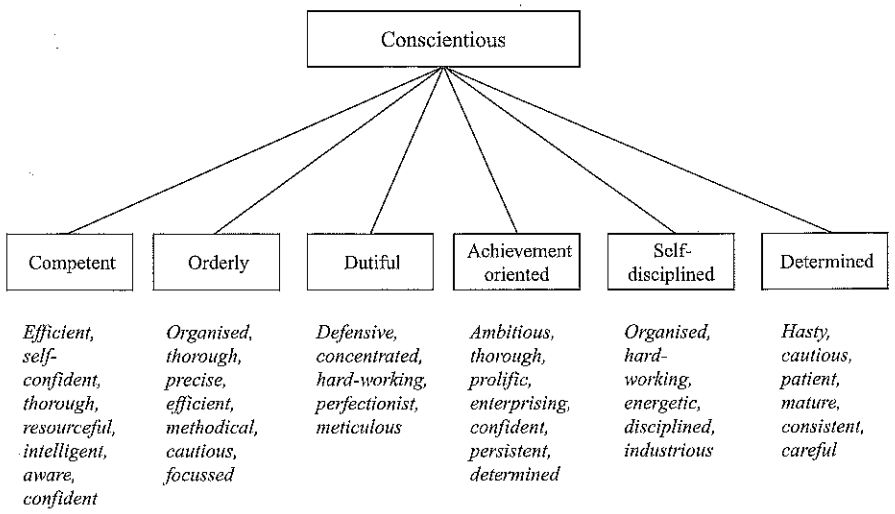


Figure 7.11 *Structure and facets of Conscientiousness (based on Costa & McCrae, 1992)*

That said, measures of Conscientiousness vary quite significantly, with some focusing more on the 'motivational' component (e.g., being ambitious, driven and proactive) and others emphasising the 'tidiness' component (e.g., being organised, clean and somewhat of a perfectionist). Moreover, high correlations between measures of Conscientiousness and Integrity inventories (see Section 7.17) suggest that another important aspect of Conscientiousness is dutifulness or moral responsibility (what Sigmund Freud called the Super-Ego). It has even been suggested that Conscientiousness is one of the main facets of integrity, though given that different integrity scales are often uncorrelated there is some uncertainty as to what integrity scales are assessing in the first place.

The first compelling evidence for the validity of Conscientiousness in the context of personnel settings was provided by Barrick and Mount (1991) (see Section 7.4). Indeed, the consistent validity of Conscientiousness was the major finding in that study. Thus the authors concluded that:

> The major finding [of their meta-analysis] was that one of the Big Five dimensions, Conscientiousness, correlated positively with job performance in all five occupational groups. Individuals who are dependable, persistent, goal directed and organised tend to be higher performers on virtually any job; viewed negatively, those who are careless, irresponsible, low achievement striving and impulsive tend to be lower performers on virtually any job. (Mount & Barrick, 1998, p. 851)

Unlike GMA, which provides a measure of people's 'can do' or maximal performance, Conscientiousness is especially important as marker of typical performance or people's 'will do'. This is especially noticeable under low extrinsic reward conditions because conscientious individuals will differ from

their non-conscientious counterparts in their tendency to strive for excellence and set themselves high targets, which they work hard to achieve. In that sense, Conscientiousness is the ideal 'partner' of GMA and provides the motivational information that GMA tests do not. This is in line with a longstanding tradition in individual differences and I/O psychology to conceptualise aptitudes and conative or motivational traits as essentially independent.

Research evidence from organisational studies suggests that Conscientiousness is negatively related to GMA (Moutafi, Furnham & Paltiel, 2004; Moutafi, Furnham & Tsaousis, 2006) and we have interpreted this association in terms of the compensational function of Conscientiousness in competitive settings. Thus, if employees are not 'equipped' with the necessary cognitive ability levels to perform well at work, they will compensate for this by becoming more conscientious and relying more on their dutifulness, self-discipline and organisational skills (Chamorro-Premuzic & Furnham, 2005). That said, most of the variance in Conscientiousness is unaccounted for by cognitive ability measures, and the effects of Conscientiousness on work-related outcomes – or indeed academic performance (Chamorro-Premuzic & Arteche, in press) – are largely independent of GMA. Thus, it is not difficult to echo Mount and Barrick's conclusion that 'no matter what job you are selecting for, if you want employees who will turn out to be good performers, you should hire those who work smarter and work harder' (Mount & Barrick, 1998, p. 856).

It is also likely that the correlation between Conscientiousness and job performance underestimates the real impact of Conscientiousness at work. First, there is a self-selective bias in any competitive setting that narrows the range in Conscientiousness. Conscientious people are not only more likely to apply for better jobs in the first place, they also have higher levels of organisational switching (the tendency to look for better jobs within and outside one's organisation) (Vinson, Connelly & Ones, 2007). Second, Conscientiousness is negatively related to almost any indicator of counterproductive behaviour, not just at work but in life in general. Indeed, a comprehensive meta-analysis reported that Conscientiousness is negatively correlated with virtually every form of unhealthy and dangerous behaviour (Bogg & Roberts, 2004).

Table 7.1 reproduced from this study summarises the findings. In particular, conscientious people are less likely to drink in excess, use drugs, drive in a risky way and use violence. Although there are ethical boundaries to recruiting on the basis of such behaviours and even asking people to report them, it is hard to conceive of many employers who, on this basis, would prefer to hire individuals who are low in Conscientiousness. The associations reported by Bogg and Roberts highlight the prophylactic nature of this personality trait in relation to counterproductive behaviours (at work and in general). In Section 6.11, we have seen how research evidence has suggested that higher GMA may provide individuals with the necessary cognitive tools to predict the negative consequences of a wide range of counterproductive behaviours (enabling them to foresee potential punishments). In a similar vein, higher Conscientiousness

Table 7.1 *Conscientiousness and health-related behaviours (from Bogg & Roberts, 2004; with permission from Roberts, APA copyright)*

Health behavior	*r*	No. of studies	*N*	95% CI Lower	95% CI Upper	*Q*
Activity	.05	17	24.259	.04	.07	136.80
Excessive alcohol use	−.25	65	32.137	−.25	−.24	1,109.89
Drug use	−.28	44	36.573	−.29	−.27	662.21
Unhealthy eating	−.13	14	6.356	−.16	−.11	126.78
Risky driving	−.25	21	10.171	−.27	−.24	422.63
Risky sex	−.13	25	12.410	−.15	−.11	76.75
Suicide	−.12	19	6.087	−.14	−.09	123.47
Tobacco use	−.14	46	46.725	−.15	−.13	352.83
Violence	.25	25	10.277	10.26	−.24	119.22

may strengthen individuals' motivation to avoid these behaviours and is probably a bigger determinant of such behaviours than is GMA: one can easily think of individuals who are bright enough to foresee prospective punishments but lack the necessary ego-strength to inhibit their behaviours. Long before the Big Five taxonomy was established, Hans Eysenck referred to such non-ability predictors of risky and anti-social behaviours in terms of Psychoticism (Eysenck & Eysenck, 1977), which, in the Big Five, is represented in terms of low Agreeableness, high Openness and low Conscientiousness (Costa & McCrae, 1995).

Further evidence for the importance of Conscientiousness in personnel selection derives from the wide range of recent studies that reported consistent positive associations between this personality trait and academic performance, a key antecedent of personnel selection and occupational performance. Several studies reported moderated (uncorrected) validities for Conscientiousness and its primary facets as predictors of various academic performance outcomes (Chamorro-Premuzic & Furnham, 2003a, 2003b, 2004), and recent meta-analytic studies have confirmed that the association of Conscientiousness and post-secondary educational achievement is robust (Noftle & Robins, 2007; O'Connor & Paunonen, 2007). Moreover, newer constructs within the Conscientiousness spectrum have been developed and validated in the context of academic performance, notably the construct of Grit (Duckworth, Peterson, Matthews & Kelly, 2007), highlighting the importance of self-control in achievement-related outcomes.

7.7 Neuroticism: it helps if you are calm

Big Five traits other than Conscientiousness have also been associated with important occupational outcomes, albeit less consistently. Such is the case

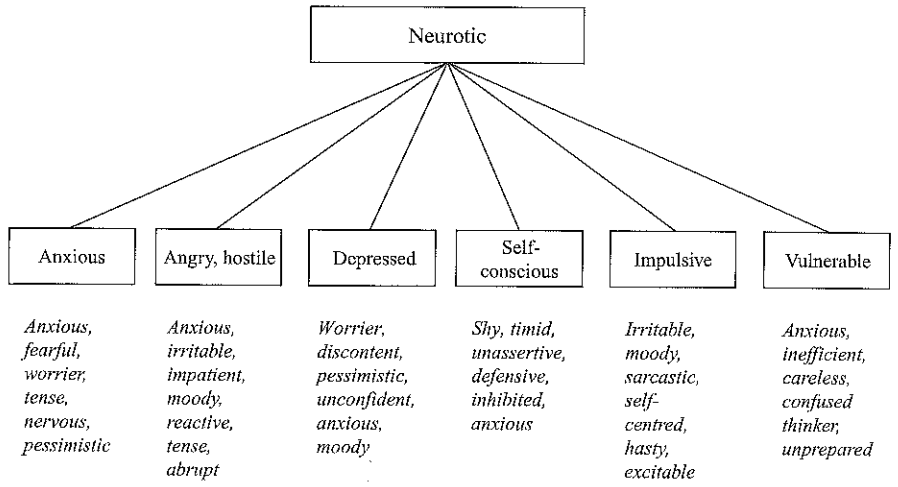


Figure 7.12 *Structure and facets of Neuroticism (based on Costa & McCrae, 1992)*

with Neuroticism (Emotional Stability), which has been linked to job and training performance particularly under stressful conditions. This is consistent with the interpretation of the Neuroticism trait (also known as trait anxiety) as a marker of individual differences in stress-reactivity, the tendency to experience low confidence and heightened negative affect, and a predisposition to anxiety in general (Chamorro-Premuzic, 2007). In the Big Five of Costa and McCrae (1995), Neuroticism comprises the subscales or primary facets of anxiety, angry hostility, depression, self-consciousness, impulsivity and vulnerability (see Figure 7.12). However, it should be noted that Neuroticism inventories assess individual differences in normal rather than clinical samples/subjects, hence they are suitable for personnel selection.

The main meta-analysis on the Big Five and work-related outcomes (reviewed above) yielded inconsistent findings with regard to the validity of Neuroticism. Thus Hough *et al.* (1990) found significant associations for this trait and different measures of military performance (i.e., effort and leadership, personal discipline and physical fitness bearing), and similar associations were found for civil professions by Tett *et al.* Likewise, Hough (1998b) found that emotionally stable recruits performed more effectively in combat, and similar findings were reported for European military samples by Salgado (1998). However, Salgado's (1997) EC meta-analysis reported only a modest correlation between Neuroticism and training success in civil occupations, and, moreover, Barrick and Mount's meta-analyses reported validities close to zero for Neuroticism.

It has even been pointed out that Neuroticism is not always negatively associated with occupational performance outcomes. Indeed, major theories of personality have long postulated that if coupled with specific ability or personality traits Neuroticism may have positive impacts on various task performances. Most

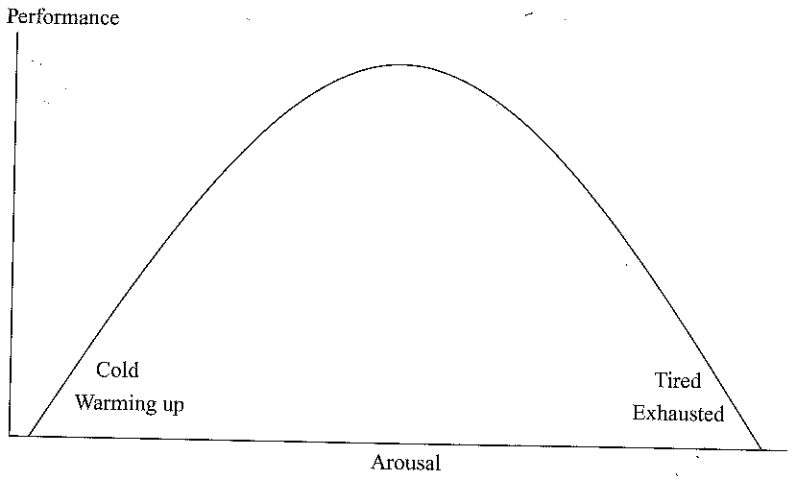


Figure 7.13 *Yerkes–Dodson law*

notably, Gray's (1988) 'reward sensitivity theory' (RST) of personality hypothesises that higher Neuroticism – combined with lower Extraversion – is linked to higher levels of learning in the presence of 'sticks' or punishment, and there is evidence in support of this hypotheses (Corr, Pickering & Gray, 1997). On the other hand, Hans J. Eysenck (1994) argued that Neuroticism and cognitive ability are linked via arousal (baseline levels of cortical arousal or brain reactivity), and Robinson (1997) conducted several EEG studies showing that higher IQ levels are linked to intermediate levels of arousal. Indeed, this idea is consistent with a well-established principle known as the Yerkes–Dodson law, which states that performance is best at an intermediate level of arousal (Yerkes & Dodson, 1908; see Figure 7.13). This explains why, under some circumstances, notably low situational pressure or tasks that are under-arousing, Neurotic individuals have an advantage over their stable counterparts because they are naturally more alert to potential environmental threats. In line with this, studies on air traffic controllers tend to report superior performance by Neurotic individuals (Matthews, 1999).

The fact that meta-analytic studies rarely explored curvilinear effects of Neuroticism on job or training performance may explain the inconsistencies of their results. Moreover, in order to understand the relationship between Neuroticism and job-related outcomes it is fundamental to take into account situational factors such as pressure, whether individuals are motivated by rewards or punishment, and even how hard the task is (for which cognitive abilities should be taken into account).

That said, the bulk of evidence suggests that if any occupational consequences of Neuroticism generalise, these refer to the negative effects of Neuroticism on work-related outcomes and, conversely, the fact that Emotional Stability enables individuals to perform as well as they could in most settings, including under elevated pressure. This idea is in line with recent studies on the

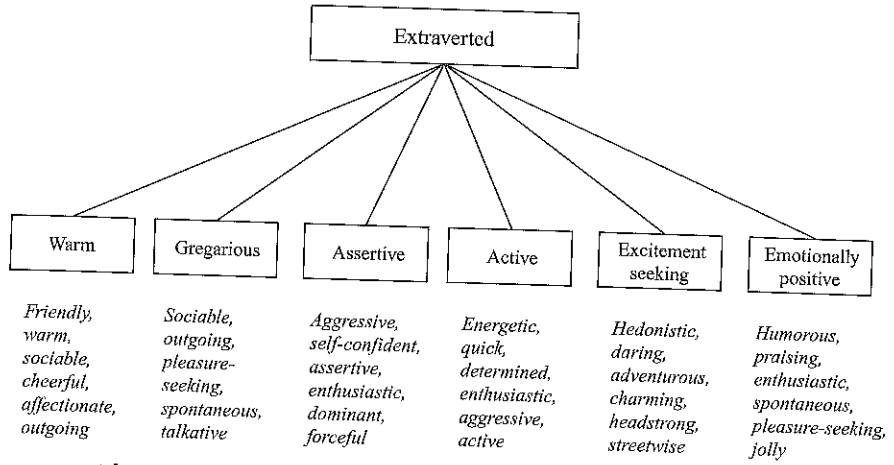


Figure 7.14 *Structure and facets of Extraversion (based on Costa & McCrae, 1992)*

personality–intelligence interface. For example, Moutafi *et al.* (2006) showed that the relationship between Neuroticism and (lower) IQ test performance was explained by test-anxiety; Chamorro-Premuzic, Ahmetoglu and Furnham (2008) showed that Neuroticism was strongly and positively linked to test-anxiety (even when controlling for other individual differences). Perhaps more importantly, Neuroticism has been recently reported to be a consistent negative predictor of post-secondary academic performance (Chamorro-Premuzic & Furnham, 2002, 2003a, 2003b; O'Connor & Paunonen, 2007). Thus even though Neurotic individuals do not differ in cognitive ability levels from their emotionally stable counterparts, they are more likely to underperform in ability tests and real-life tasks, especially in the presence of pressure (Chamorro-Premuzic & Furnham, 2004) (see again Figure 7.1). Recent work on emotional intelligence (see Sections 7.21, 7.22, 7.23), a strong negative correlate of Neuroticism, adds support to the idea that being emotionally stable or in control of one's emotions is beneficial in the workplace.

7.8 Extraversion: positive, sociable people-people

Another trait that has been associated with job-related outcomes, albeit inconsistently, is Extraversion. In the Big Five model of Costa and McCrae (1995) (but see Hogan, 1986, and Hogan & Holland, 2003, for a different view) Extraversion includes the subfacets or primary traits of warmth, gregariousness, assertiveness, activity, excitement seeking and positive affect/emotionality (see Figure 7.14). Thus positive associations between Extraversion measures and performance in managerial and sales jobs have been found and interpreted in terms

Table 7.2 *Task-dependent correlates of Extraversion (based on Matthews, 1999)*

Extraversion level	High	Low
Divided attention	+	−
Long-term memory	−	+
Reflective problem solving	−	+
Resistance to distraction	+	−
Retrieval from memory	+	−
Short-term memory	+	−
Vigilance	−	+

of the higher interpersonal competence of extraverts, something that also applies to more agreeable employees (Mount, Barrick & Stewart, 1998).

Extraversion seems particularly important when 'coupled' with Emotional Stability (low Neuroticism), what some people refer to as the *happy personality* (Chamorro-Premuzic, Bennett & Furnham, 2007). This idea is congruent with two major (pre-Big Five) approaches to personality, namely Hans Eysenck's and Jeffrey Gray's personality models. Eysenck was instrumental in showing that stable Extraverts are better adapted and happier than their introverted counterparts, whereas Gray argued that impulsivity is characterised by a combination of high Extraversion and high Neuroticism. Along these lines, a recent study by Judge and colleagues found that 'happier' employees (those high in Extraversion *and* low in Neuroticism) performed significantly better than extraverted or stable employees alone (Judge & Erez, 2007). Likewise, previous studies found that Neuroticism and Introversion combine to have deleterious effects on performance. For example, a study on fighter jet pilots found that a quick (ten-minute) measure of Extraversion and Neuroticism predicted training success and failure (again, being Extraverted and Stable was linked to success) (Bartram & Dale, 1982). These finding are important because they highlight interactive effects underscored by regression-based studies focussed on the individual contribution of each predictor. Moreover, the common assumption that the Big Five are truly orthogonal (unrelated) underscores potential synergistic effects that may result from the overlap between them, notably Extraversion and Neuroticism (which tend to be negatively correlated) (Chamorro-Premuzic, 2007). Indeed, Judge and colleagues also found that the intersection between Extraversion and Neuroticism predicted job performance better than the unique aspects of these traits.

Unlike Costa and McCrae (1995), Hogan (1986) interpreted the dimensions of Extraversion in terms of sociability, ambition, adjustment, likeability, prudence (lack of) and intellect. The main difference between this model and the one proposed by Costa and McCrae is that Hogan splits Extraversion into the two major subcategories of Sociability and Ambition. Thus some Extraverts may be characterised more by their tendency to experience positive affect, be sociable

and enjoy the company of others (as well as dread being alone), whilst in other Extraverts the main trait would be dominance, self-confidence and leadership. As seen, some of these various facets of Extraversion have been conceptualised by Costa and McCrae (1995). However, the issue is whether they really load onto one single higher-order factor or represent two distinct traits (as proposed by Hogan). This is an important question because some aspects of Extraversion may be more beneficial at work than others. Indeed, one may think of jobs that suit 'Sociable' extraverts more than 'Ambitious' extraverts and vice versa.

Last, but not least, the inconsistent associations between Extraversion and job-related outcomes have also been explained in terms of the cognitive characteristics of the Extraversion trait. Specifically, the cognitive nature of the task will determine whether extraverts or introverts are more likely to perform better (Matthews, 1999). Thus extraverts would have an advantage over introverts in tasks that require divided attention (e.g., writing while listening to music, or reading while watching TV) because of their lower levels of distractibility. This is consistent with Eysenck's arousal theory of Extraversion/Introversion. Moreover, Extraverts would also have an advantage in tasks requiring retrieval from short-term memory (and better memory retrieval in general). However, if tasks require vigilance introverts have the advantage over extraverts as the latter would try to compensate for their lower levels of cortical arousals by attending to task-irrelevant stimuli. Tasks that require long-term memory retrieval or problem solving (where accuracy matters more than speed) also benefit Introverts more than Extraverts.

7.9 Agreeableness: getting along, caring and sharing

Agreeableness seems to be advantageous in jobs requiring interpersonal interactions or where *getting along* is paramount (Hogan, Rybicki, Motowildo & Borman, 1998; Mount *et al.*, 1998). A typical case is customer service jobs, and indeed Agreeableness has been found to predict performance on these jobs quite well (Hurtz & Donovan, 2000), especially if based on teamwork rather than individualistic tasks (Barrick, Stewart, Neubert & Mount, 1998). Agreeableness also seems to moderate the effects of Conscientiousness – the strongest personality trait correlate of job performance – or work-related outcomes (Witt & Ferris, 2003). Thus people who are Conscientious but Disagreeable will tend to have conflicts with others, whereas people who are Conscientious and Agreeable will benefit from the synergistic effects of discipline and cooperation.

That said, cooperation is not always a driver of productivity and in some cases may hinder personal success. Thus Hogan and colleagues have argued that individuals higher in Agreeableness are less likely to be driven by social status, particularly if they have to attain it at the expense of others. In situations where collaboration and competition are mutually exclusive, Agreeableness may be

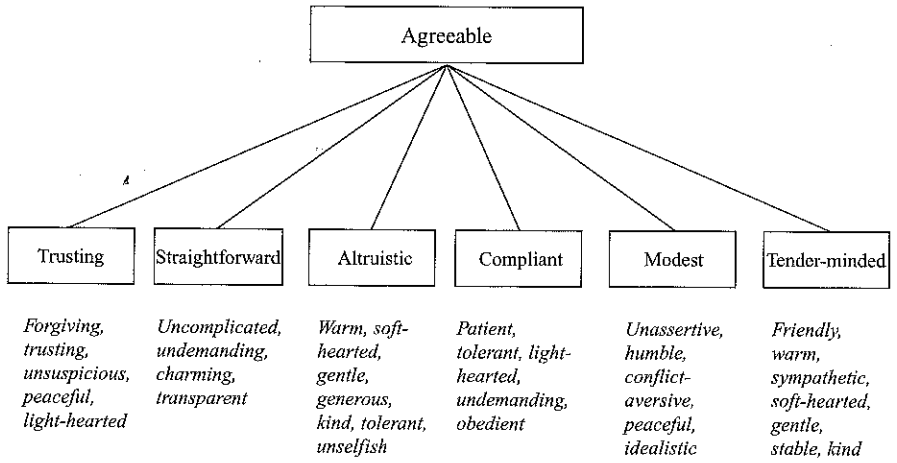


Figure 7.15 *Structure and facets of Agreeableness (based on Costa & McCrae, 1992)*

negatively linked to job performance as disagreeable employees may prioritise *getting ahead* over getting along. Interestingly, and in line with Hogan's idea, most of the empirical evidence suggests that Agreeableness is *negatively* rather than positively linked to career success (operationalized in terms of income). Most notably, two large meta-analytic studies reported associations of −.24 (US) and −.11 (Europe) for Agreeableness and income (Boudreau, Boswell & Judge, 2001), and a smaller but longitudinal study reported a coefficient of −.32 for these variables (Judge, Higgins, Thoresen & Barrick, 1999). This is consistent with the idea that high achievers, and especially leaders, may be characterised by higher levels of Machiavellianism, a trait that assesses individuals' willingness to take advantage of others in order to accomplish their own goals (Austin, Farrelly, Black & Moore, 2007; Drory, 1980).

7.10 Openness: intellectual, imaginative, artistic jobs

The last of the Big Five personality traits is Openness to New Experiences, often simply referred to as 'Factor V', not least because it is unclear what this trait really encompasses. In Costa and McCrae's (1995) model, high Openness scorers are characterised as being driven by fantasy, artistic, feeling-oriented, action-oriented, ideational and having liberal values (see Figure 7.16). Other Big Five taxonomies emphasise the higher levels of need for cognition in open individuals. Openness is correlated with knowledge-based or 'crystallised' aspects of intellectual ability (Ackerman & Heggestad, 1997; Chamorro-Premuzic & Furnham, 2005); artistic professions require higher levels of Openness, an association that has been interpreted in terms of the higher likelihood of Open individuals to 'invest' in intellectual activities and acquire more knowledge

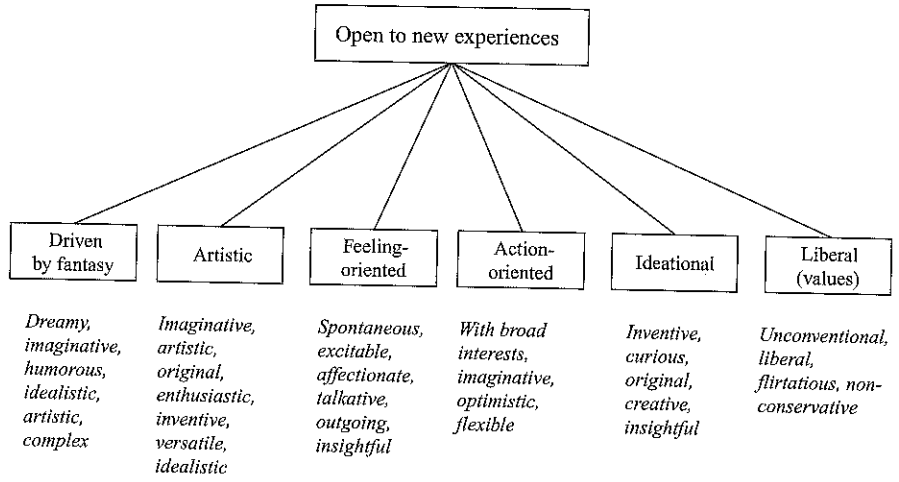


Figure 7.16 *Structure and facets of Openness (based on Costa & McCrae, 1992)*

(Cattell, 1950; Chamorro-Premuzic & Furnham, 2006a). However, Openness has been inconsistently related to work performance and seems relevant only in artistic or creative jobs, though it does affect training success.

7.11 It's not all about performance: validity of the Big Five as predictors of job satisfaction

Although this chapter is primarily concerned with the question of whether personality traits and inventories are useful predictors of work and training performance, performance is not the only thing that matters. Indeed, an important work-related criterion (which may actually drive performance levels itself) is job satisfaction, and psychologists have long argued that individual differences in job satisfaction or the extent to which employees are happy at work are largely influenced by dispositional or personality factors. Moreover, industrialised economies have come to realise in recent times that 'getting richer' does not mean 'getting happier', and in the quest for the right work–life balance an important component is the degree of gratification that people get from their work, and this seems to vary largely independently of performance or financial success.

What personality characteristics determine individual differences in job satisfaction? It seems that emotional adjustment plays the most important role. Almost seventy-five years ago, Hoppock assessed the level of job satisfaction of psychologists and found that those higher in emotional adjustment were more satisfied at work (Hoppock, 1937). Since then a large body of studies has reported consistent associations between affective dispositions (e.g., Neuroticism, emotional intelligence and Extraversion) and various indicators of job satisfaction. However,

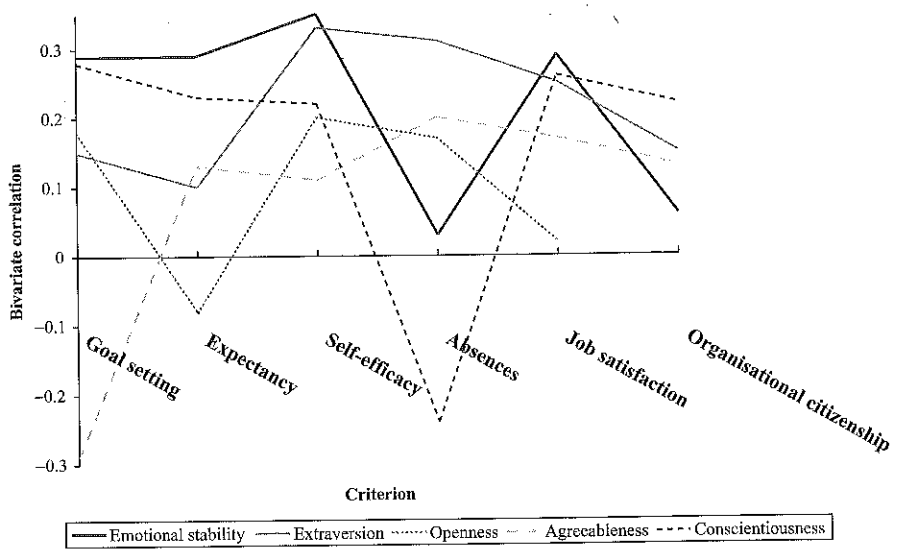


Figure 7.17 *Importance of the Big Five as predictors of motivational outcomes*

until recent meta-analyses made sense of the literature it was difficult to identify overall trends in the data and quantify the impact of major personality dimensions on job satisfaction and related outcomes.

Figure 7.17 (based on Judge & Ilies, 2002; Judge, Heller & Mount, 2002; Judge, 1997; Organ & Ryan, 1995) summarises the meta-analytic evidence for the validity of the Big Five personality traits as predictors of these various motivational outcomes, namely *goal setting*, *expectancy*, *self-efficacy*, *absences* (an indicator of dissatisfaction), *job satisfaction* and *organisational citizenship*. Each trait is represented by a different shade and correlations for different criteria are plotted along the *x* axis. For instance, the correlation between Openness and goal setting was .18, whilst the correlation between Conscientiousness and absences was −.24 (negative correlations are plotted below the *x* axis and vice versa).

As shown, all five factors contributed significantly to the prediction of individual differences in these outcomes, though validities varied according to the outcome and predictor (from −.08 to .35). For example, Openness made a negligible contribution to the prediction of job satisfaction and a very modest contribution to explaining expectancy motivation. Thus whether you are Open or not, your levels of job satisfaction and expectancy motivation remain pretty much the same. On the other hand, Emotional Stability contributed to the prediction of most outcomes but was only weakly related to absences and organisational citizenship. Extraversion and Agreeableness were linked to all outcomes, though sometimes inconsistently: for example, both traits were linked to higher absenteeism rates and disagreeable people had higher levels of goal-setting motivation than their agreeable counterparts did. Thus the only consistent personality predictor of the

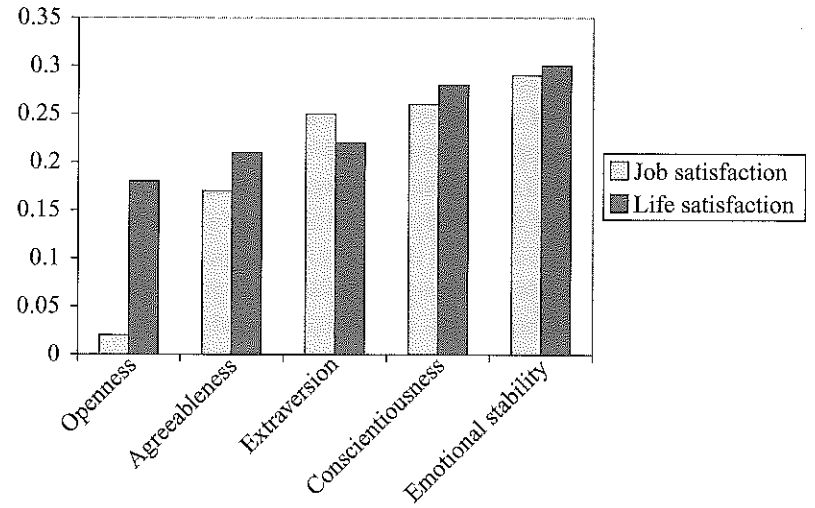


Figure 7.18 *Big Five as predictors of job and life satisfaction*

outcomes examined was Conscientiousness. In fact the validities for Conscientiousness were pretty much equal across all criteria examined.

Research evidence also indicates that satisfaction and dissatisfaction are not two opposite extremes of the same dimension (though see Yik, Russell & Barrett, 1999), but, rather, two distinctive factors. Thus employees may not be satisfied without, however, being dissatisfied, and, likewise, someone may not dissatisfied without, nonetheless, being satisfied (Herzberg, 1965). This is consistent with a well-established tradition in psychology to distinguish between 'positive' and 'negative' affect as two separate dimensions of emotionality. Individuals high in positive affect (who tend to be stable and extraverted) are naturally predisposed to experience joy, excitement and enthusiasm, whereas those high in negative affects have a tendency to experience guilt, anger, fear and concerns (Watson, Clark & Tellegen, 1988). Likewise, Jeffrey Gray's neuropsychological theory of personality postulates the independence of two neural circuits responsible for learning and motivation, namely the behavioural inhibition system (BIS) and behavioural approach system (BAS), which are linked to negative and positive affect, respectively (Gray, 1982). However, for the purposes of predicting job satisfaction, the distinction between positive and negative affect as different dimensions seems overrated, as both factors are linked to job satisfaction measures. For example, a meta-analysis (Connolly & Viswesvaran, 2000) reported that job satisfaction correlated positively with positive affect measures (.49) and negatively with negative affect measures (−.33).

The best meta-analytic estimate of the dispositional basis of job satisfaction was provided by Judge *et al.* (2002) (see Figure 7.18). As shown, Conscientiousness, Emotional Stability and Extraversion had validities larger than .20 in the prediction of job satisfaction, whereas Agreeableness was more

weakly but still significantly associated with this outcome (and Openness was unrelated). Interestingly, looking at the correlations between personality and *life satisfaction* – estimates are taken from DeNeve and Cooper (1998) – reveals a very similar pattern of results (the only exception is for open individuals, who seem to be much happier than their less open counterparts in life, but not at work!). Thus dispositional sources underlying individual differences in the extent to which people are happy at work operate in a similar way *beyond the workplace*. This is important because it shows that the scope of personality is not limited to people's jobs, and that selecting employees with a higher dispositional basis for satisfaction (i.e., those emotionally stable, extroverted and conscientious, and to a lesser extent agreeable) will result in employees who are happier in both the job and in general. The authors reported an overall correlation of .41 for all Big Five combined and job satisfaction, which shows that the five factors are good indicators of a person's likelihood to be happy at work.

7.12 Criticisms to the Big Five (and personality inventories in general)

Even since personality inventories were developed there have been objections to the use of such tests in personnel selection (and, in the wider sense, for any decision-making process). In the context of work psychology the two main criticisms are that it is easy to fake responses to a personality inventory and that personality traits are only weak predictors of occupational outcomes. In the forthcoming sections we discuss these two criticisms.

7.13 Faking

Several arguments have been put forward in defence of personality inventories attempting to persuade researchers and practitioners that faking is *not* really problematic in the context of personality assessment. These go as follows:

1. Since applicants are likely to fake good but not to fake bad, the validity of personality inventories is only threatened by faking at the high-end of the distribution (Mueller-Hanson, Heggestad & Thorton, 2003). This means that whereas some candidates will inevitably manage to 'crack the test' and obtain desirable scores simply because they lied and were able to identify the better responses, it is unlikely that candidates with unfavourable profiles would have faked the results. Accordingly, it has been suggested that self-report inventories are used for filtering-out rather than selecting-in candidates.
2. Studies in which faking was detected showed that personality traits still predicted job-relevant outcomes (Hough *et al.*, 1990; Ones, Viswesvaran & Reiss, 1996).

3. Personality inventories administered during high-stake situations still predict job performance and other work-related outcomes (Ones & Viswesvaran, 2003).
4. Laboratory studies, where participants (especially students) are instructed to 'fake good', yield findings that do not necessarily generalise to 'high-stake' settings (see point 1 below) (Viswesvaran & Ones, 1999).
5. The importance of social desirability (as measured by meta-scales that detect extremely positive responses) appears to have been over-rated. First, social desirability scales fail to predict desirable work-outcomes (Viswesvaran, Ones & Hough, 2001). Second, controlling for social desirability has no impact on the validity of personality traits, suggesting that the effects of personality traits on job-related outcomes cannot be accounted for by social desirability scales (Ones *et al.*, 1996). Third, social desirability does not affect the structure of personality inventories (Ellingson, Smith & Sackett, 2001).
6. Being able to fake is a sign of social adjustment in itself (Hogan, 2005). As Hogan pointed out, people's responses to personality items can be interpreted merely as an 'automatic and often non-conscious effort to negotiate an identity with the interviewer/test developer' (Hogan, 1992, p. 902). This is an important point which even critics of personality inventories endorse. Clearly, if you can pretend to be good in a test you can probably do the same in the workplace (and life in general). As summarised neatly by Murphy, 'if faking is defined in terms of saying what you think you ought to say rather than what you want to say that is called civilization' (in Morgeson *et al.*, 2007, p. 712). Social desirability is especially important when it comes to performance assessed by supervisors rather than objective measures. Moreover, in some jobs (highly interpersonal ones, such as customer service and relations, are the obvious ones) the ability to fake good may be essential. Dipboye points out (in Morgeson *et al.*, 2007) that if you are working in Disney you would probably be expected to fake good – and even happiness – most of the time, and the same may apply to McDonalds, at least if you want to make it to employee of the month. Moreover, individual differences in social desirability have been explained in terms of Conscientiousness, Emotional Stability and Agreeableness, and these traits predict performance in the first place.

On the other hand, the most frequently discussed arguments against the validity of personality inventories in personnel selection have highlighted that fact that faking *is* a problem:

1. Several studies have shown that job applicants tend to score 'higher' (i.e., show more favourable or desirable profiles) than incumbents, which can be attributed to the formers' motivation to fake good on these scales (Barrick & Mount, 1996; Hough, 1995; Rosse, Stecher, Miller & Levin, 1998).
2. Laboratory studies, such as those instructing participants to 'fake good', show that it is very easy to identify the desirable or correct responses to a personality inventory (Viswesvaran & Ones, 1999) (though see point 4 above).

Can people fake? → almost 100% of the studies (39) suggest YES
Do they fake? → at least 50% of the studies (14) suggest YES
Do some fake more than others? → 100% of studies (3) suggest YES
Does faking depend on the situation? → at least 70% of studies (7) suggest YES
Are self-deception and impression management different factors? → 80% (4) YES
Does faking harm the validity of personality scales? → 50% (18) YES*
Does faking affect the structure of personality? → 50% (4 studies) YES
Can faking be mitigated? → 30-40% of studies (10) suggest YES
Can faking be detected? → 40% of studies (33) suggest YES

*This point is rejected by one of the authors of the same article, Hollenbeck, who correctly points out that the decrease in validity attributed to faking in those 50% of studies is at most marginal.

Figure 7.19 *Review of faking (by Michael Campion, in Morgeson* et al., *2007)*

3. Even if the structure and overall validity of personality inventories may not be affected by faking, faking produces changes in the rank order of candidates – meaning faking will influence hiring decisions – because some may fake more than others, i.e., there are both personality and situational determinants of faking not captured by personality scores (Mueller-Hanson, Heggestad & Thornton, 2006).
4. Even if social desirability scales do not correlate with work-related outcomes, and even if they do not explain the validity of personality traits predicting those outcomes, faking may still be an issue simply because social desirability scales may fail to successfully measure faking in the first place (Ellingson, Sackett & Hough, 1999).

Some of the discrepancies noted above can be attributed to the lack of compelling evidence either against or in support of the idea that faking poses a problem. In an attempt to provide a comprehensive review of the literature, Michael Campion (in Morgeson *et al.*, 2007) examined the salient studies on faking (see Figure 7.19), concluding that:

> Four overall conclusions can be drawn from this review of the research literature on faking in personality tests. First, the total number of studies on the topic is large, suggesting that faking has been viewed as an important problem. Second, people can and apparently do fake their responses on personality tests. Third, almost half the studies where criterion-related validity was studied found some effect of faking on criterion-related validity. Fourth, there has been substantial research devoted to techniques for detecting and mitigating faking, but no techniques appear to solve the problem adequately. (p. 691)

7.14 How to overcome the problem of faking

Several suggestions have been put forward to overcome the problem of faking (by those who see faking as a problem). These range from simply warning the test-takers (Vasilopoulos, Cucina & McElreath, 2005) to comparing the structure of personality inventories in candidates and low-stake samples (Kuncel & Borneman, 2007). The latter requires test-interpreters to examine whether items that are normally uncorrelated are positively correlated in the applicant/employee/candidate. Others have argued that forced-choice scales, such as asking people whether they are X *or* Y (rather than how representative X is), reduce the problem of faking (Villanova, Bernardin, Johnson & Dahmus, 1994). On the other hand, there is the option of *other-* rather than self-reports, which have been shown to explain additional variance over and above self-reports in the context of occupational (Barrick, Mount & Strauss, 1994), and even secondary school performance (Bratko, Chamorro-Premuzic & Saks, 2006; see also Murphy's comments in Morgeson *et al.*, 2007).

Yet there are issues with all of the recommendations proposed. First, warnings would probably not have the same impact on every test-taker, meaning there are surely individual differences in responses to faking-related warning, which will deter only some of the potential fakers. Second, contrasting the structure of individual responses with previously collected data (from low-stake contexts) may be generally effective but harm honest respondents with an individual or different-from-average-respondents profile. Third, it is not necessarily the case that forced-choice items will prevent 'street-wise' fakers from spotting the most desirable response. For example, a forced-choice item on Conscientiousness may ask candidates to choose between 'Hard-working, reliable and committed' on one hand, and 'Lazy, unreliable and uncommitted' on the other (this is no doubt a somewhat crude example but it does show how forced-choice items are not useful *per se*). In fact, the less obvious forced items are, the more they may compel participants to choose between two options that may both apply to them, e.g., 'Organised, methodical and tidy' versus 'Creative, spontaneous and intuitive' (which may assume that aspects of Conscientiousness and creativity are negatively correlated). Fourth, even if other-ratings have incremental validity over self-reports, it is not easy to collect these data, especially in the context of personnel selection (if recruiting from outside the company). Moreover, it has been pointed out that although the accuracy of other-reports seems to depend on the level of familiarity between the observer and the candidate, familiarity levels may be positively correlated with the observers' motivation to lie and indeed fake good about the candidate (Connolly, Kavanagh & Viswesvaran, 2007). Besides, self- and other-ratings are probably not measuring exactly the same construct.

A somewhat idealistic yet commonsense proposition argued that the way to overcome the problem of faking is to shift from an 'interrogational' or at least inquisitive approach to personality assessment to a more participative, interactive,

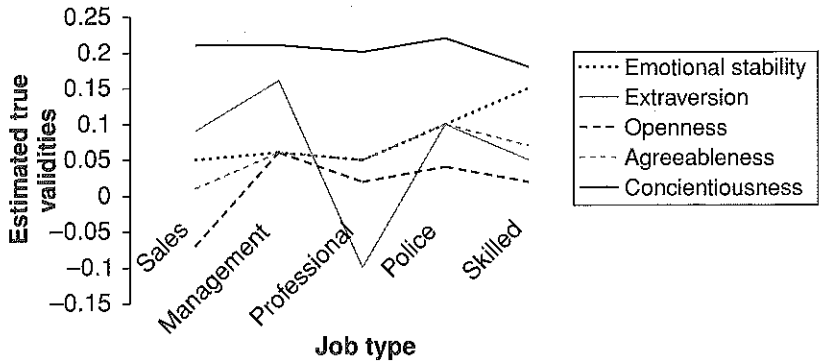


Figure 7.20 *Meta-meta analysis of the Big Five and job performance*

collaborative process where information is not withdrawn from the candidates and where the constructs being assessed are clearly specified to the candidates. Assuming that they will understand the importance of 'fitting' into the right organisation and job (see Section 7.26), they may be expected to respond honestly if this is done (see Dipboye's comments in Morgeson *et al.*, 2007). That said, given that organisations tend to seek the same type of candidates (in personality terms, highly conscientious, stable and perhaps extraverted) it would be impossible for everyone to do well in personnel selection unless they decided to fake good in the expected direction of the employers/recruiters.

7.15 Low validity

Even if faking can be overcome, or in cases where it does not seriously threaten the validity of personality traits as predictors of work-related outcomes, critics of the use of personality inventories in personnel selection have another, often more fundamental, objection, namely the fact that the magnitude of the association between personality traits and the predicted criteria is modest at best, and often non-significant (Morgeson *et al.*, 2007). In fact, meta-analytic estimates that correct for unreliability and range restriction may overestimate the utility of personality traits in the prediction of work and training performance.

The irony is that opposite conclusions are often drawn from exactly the same data. For example, when the first quantitative reviews on the subject appeared in the 1960s researchers' recommendations against the use of personality inventories for work selection was based on an estimated validity of .09 (Guion & Gottier, 1965), whereas Barrick and Mount (1991) recommended using personality inventories based on an overall (corrected) validity of .13 for the Big Five. Although even an optimistic estimate of the validity of personality traits (adjusting and correcting for all possible drawbacks and combining all relevant traits) would hardly account for 15 per cent of the variance in job performance (Murphy,

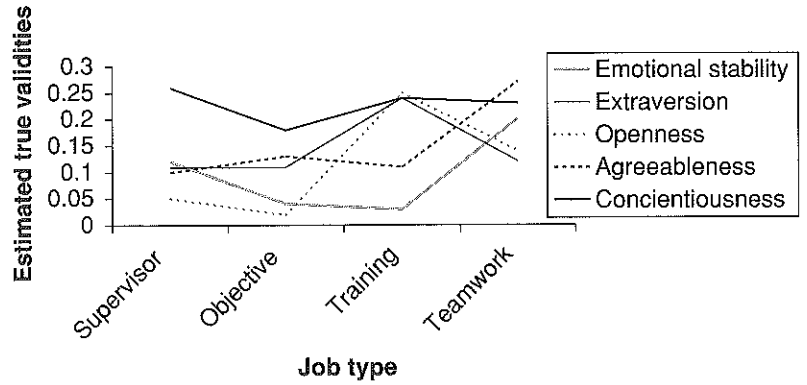


Figure 7.21 *Meta-meta analysis of the Big Five and different job outcomes*

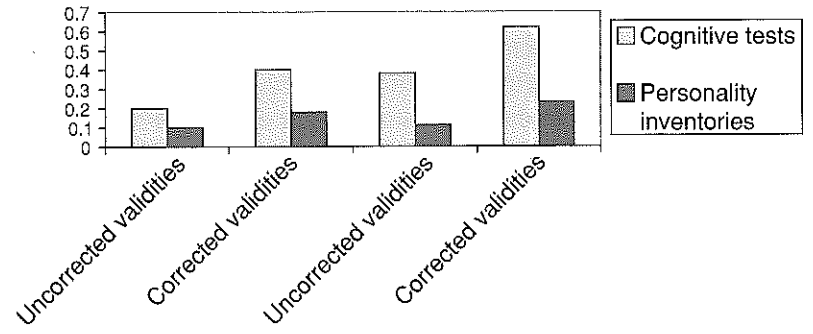


Figure 7.22 *Meta-meta-analytic estimates of the validities of cognitive ability and personality scales*

in Morgeson *et al.*, 2007), utility analyses suggested that this is an acceptable magnitude and supported the inclusion of personality inventories on the basis of the same estimate (Schmidt, Hunter, McKenzie & Muldrow, 1979).

On the other hand, subsequent meta-analyses, particularly in the 1990s (after the consolidation of the Big Five taxonomy), reported validities in the region of .40, that is, 100 per cent larger, though mainly for higher-order or 'compound' traits (Hogan, 2005; Ones, Viswesvaran & Dilchert, 2005) (see Section 7.17).

Figure 7.22 depicts the comparative predictive validity of cognitive ability tests and personality inventories in the prediction of occupational outcomes as estimated by thirteen meta-analyses in the case of cognitive ability tests and twelve meta-analyses in the case of personality inventories. As seen, the comparative validity of personality inventories (in relation to cognitive ability tests) is low, especially in the case of training criteria. However, these figures underestimate the impact of specific personality traits – like Extraversion and Emotional Stability – which may affect performance under certain settings only (unlike Conscientiousness). Thus some traits will work in some situations but not in others (which is hardly true for GMA, as seen in Chapter 6).

7.16 Specific criteria

Regardless of the magnitude of the correlation between personality scores and work-related outcomes, it is clear that the validity of personality inventories is largely dependent on the type of criterion we chose to predict. Thus, unlike with GMA, many factors moderate the effects of personality on job and training performance (though less so for job satisfaction). Of course, it should also be emphasised that the validity of 'personality' depends on what personality trait we are assessing – for example, validities are consistently higher for Conscientiousness than for Agreeableness, though in some cases (jobs low in responsibility but high in interpersonal contact) one may expect Agreeableness to be a better driver of performance than Conscientiousness. Different instruments may also be assessing different constructs, or at least different aspects of the same construct. Most notably, the extent to which 'Factor V' assesses individual differences in Openness to New Experiences, intellect, autonomy or creativity depends largely on the scale used to assess this trait. This reflects the fact that although there is now a large consensus about the explanatory power of the five-factor model, strong preferences for one inventory rather than others remain, and there are debates about how to capture the essence of some of the Big Five, notably Conscientiousness and Openness. Indeed, some of the meta-analytic studies corrected for 'construct invalidity' (Salgado, 1997) on that basis, though this procedure has been shown to overestimate the validity of personality traits (Hurtz & Donovan, 2000).

But some of the moderating factors apply equally to all of the traits. For example, there have been suggestions that the validity of personality traits *increases* with time spent on the job or in the organisation, which is consistent with the idea, albeit contested, that the validity of GMA would decrease over time (Murphy, 1989). Thus maximal performance would be important especially during the transitional phase, at early stages of the job (when motivation is high and more learning is required), but, during the maintenance phase or once employees have acquired the necessary knowledge to perform their job well, the key issue is whether they *want* to perform it or how well they want to do it. For some traits, however, the opposite pattern occurs. For instance, recent evidence suggests that the validity of self-monitoring decreases after the initial 'honeymoon' period has passed (Moser & Galais, 2007).

It is plausible to predict that the validities of personality traits will increase substantially if the correct criteria (or predictors) are chosen. As suggested by Hollenbeck's research, Openness may always be related to performance if performance requires non-conventional methods; Extraversion and Agreeableness may be consistently related to performance if performance is determined by communicational factors and collaboration, respectively, and Neuroticism may be a strong predictor of performance if performance is measured under stressful situations (see Morgeson *et al.*, 2007). To this extent, it has been suggested that practitioners, employers and recruiters design purpose-built inventories and pay particular attention to the behavioural outcomes they wish to predict. Failing that, the best way to detect faking is probably by including 'bogus' items referring to non-existent things (e.g., 'I know how to program with all versions of the Shorkipu software' or 'I have used Mokipu before').

7.17 Integrity inventories

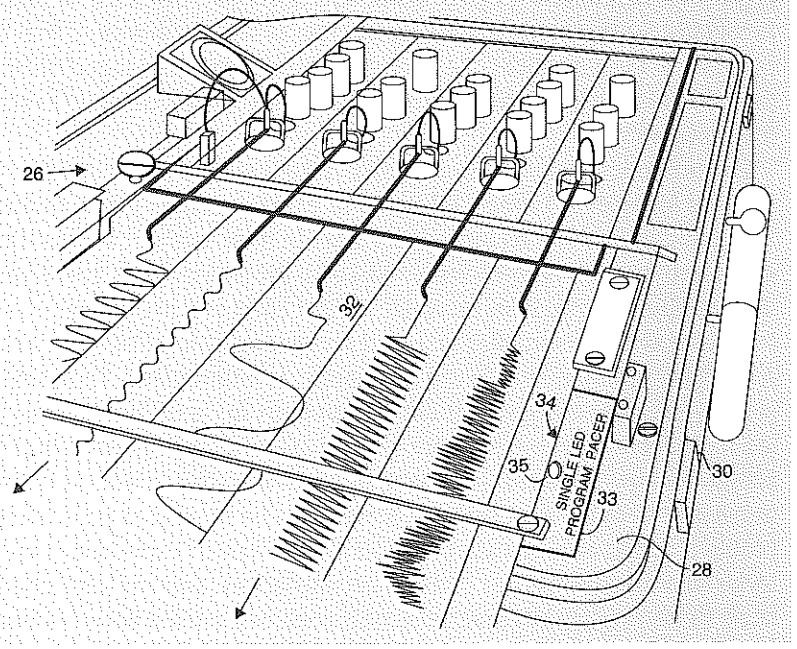
The first thing to say about integrity 'tests' is that they are usually inventories rather than tests (hence the title of this section). The second is that in the past five to ten years they have enjoyed a level of popularity in personnel selection matched only by the Big Five (hence integrity is the first non-Big Five personality trait we review in this chapter). Needless to say, the popularity of integrity inventories is justified by the amount of empirical evidence in support of their validity in the prediction of important job outcomes. Most famously, Schmidt and Hunter's (1998) seminal meta-analysis concluded that the optimal combination of psychological measures in personnel selection consisted of GMA tests plus integrity inventories; but let us examine the concept and measurement of integrity in more detail.

Box 7.3 The polygraph and the quest for an objective personality test

After many decades of intensive use in employment, in 1988, the polygraph or 'lie detector' was banned in the US for the purpose of pre-employment screening under the Employee Polygraph Protection Act (EPPA). Until then, the polygraph was a widely used and trusted tool for spotting dishonest applicants despite the fact that there had been little evidence in support of its validity. Furthermore, a growing number of studies indicated that the number of 'false positives', i.e., honest subjects who are judged dishonest on the basis of the polygraph, was too high (EPPA, 1988). So, how exactly does the polygraph work?

The polygraph provides an objective measure of various physiological markers for anxiety, namely increases in blood pressure, pulse, respiration and skin conductivity. As individuals differ in their levels of arousal (as expressed through each of these markers), the polygraph is based on *within*-subject rather than between-subject comparisons, taking baseline measures of these variables first, and comparing arousal levels at baseline with those measured while the respondent is interrogated. Thus, unlike self-reports of honesty, which ask respondents to report how honest they are in relation to various scenarios or behaviours, participants tested via the polygraph may be asked to answer any question (the content is quite irrelevant) because their honesty is meant to be measured objectively. The problem is that even honest participants are likely to experience heightened levels of arousal when put in these circumstances, and any such changes are likely to be wrongly interpreted as dishonest responding when in fact arousal was triggered by increases in anxiety that are not related to lying.

Although the polygraph has been discredited to the extent of being considered a ludicrous proposition for assessing dishonesty and honesty, it reflects a long psychological quest for the objective assessment of personality traits. Indeed, 'objective personality tests' have been researched for almost a century (Achilles & Achilles, 1917; Ball, 1929; Cattell, 1950), especially by R. B. Cattell. In some cases, such as with the objective measurement of certain primary traits like tolerance for ambiguity or boredom, objective measures of personality (e.g., looking time, response latencies and kinesis) have shown some degree of success (Messick & Hills, 1960). In recent years, there has been a renaissance of objective measures of personality no doubt thanks to the technological improvements in software and computer developments. For example, the Hyperkinetic Syndrome Assessment Method is a software-based test for measuring children's work styles, adaptation to difficulty and adjustment to feedback, and it has shown good reliability and adequate validity in educational, health and clinical settings (Proyer & Häusler, 2007). Moreover, there has also been a recent and relatively successful rediscovery of some of Cattell's objective personality scales, with suggestions that such tests are re-employed in the context of personnel selection (Santacreu, Rubio & Hernández, 2006). However, objective tests that succeed at measuring broad personality dimensions (such as the Big Five traits) or compound traits (such as Integrity) have yet to be discovered (see Sections 7.21, 7.22 and 7.23 for a discussion of objective emotional intelligence measures), and the reliability of objective measures of personality traits is poor compared to that of subjective self-reports.



Integrity inventories – based on self-reports – fall into two different categories, namely personality-like and overt (Sackett, Burris & Callahan, 1989). Personality scales for assessing integrity function in the same way as the Big Five or indeed most personality inventories, namely by including various self-report items that tap into an underlying latent dimension of integrity (or indeed several subscales, such as dependability, conformity and risk-taking, which are conceptualised as aspects of integrity). Examples of such items would be 'I am always driven by principles' or 'Even when temptations are strong I manage to exert a great degree of self-control.' The advantage of these scales is that although they rely on self-reported information it is not completely obvious to the respondent what trait or construct is being assessed. Another advantage of these scales is that they assess a broad construct that is likely to be a marker of individual differences across a wide range of contexts.

On the other hand, overt integrity scales include two types of items (Dalton & Metzger, 1993; Dalton, Metzger & Wimbush, 1994). The first type is those items that require the respondent to provide specific information on his or her past wrongdoings. Examples of such items would be 'How many times in your previous employment history have you stolen something from your employer?' or 'Have you ever pretended to be sick in order to miss a day of work?' As can be seen, these items are straight-to-the-point, but it is also easy to identify the direction of the most desirable response. Moreover, honest individuals are probably more likely to 'confess' any behaviours with negative connotations than their less honest counterparts, in which case integrity scales would end up punishing honest rather than dishonest respondents (Karren & Zacharias, 2007).

In some cases so-called 'lie scales' have been introduced to spot unrealistically desirable responses. For example, if the question were 'Have you ever lied to your employer, even on something completely trivial?' and respondents answered 'never', we would probably wonder if we are either in the presence of the most honest person on earth or if the respondent is reporting incorrect information (voluntarily or not). The second type of items for overt integrity scales uses a more impressionistic approach, which is to assess respondents' attitudes towards integrity-related behaviours, though not theirs. For example, respondents may be asked to indicate the extent to which they agree with the following statements: 'Anybody who steels something is a thief and should be punished accordingly' or 'More often than not, lies can be justified and are less damaging than the truth.' Clearly, what is being assessed with such items is people's opinions or attitudes, and the assumption is that there is a link between people's acceptance of others' dishonesty (or, to rephrase, the extent to which people regard various behaviours as morally unacceptable) and their own dishonesty (Murphy, 2000). However, research has shown that more and less honest individuals do not really differ in their attitudes towards morally acceptable or unacceptable behaviours (Nicol & Paunonen, 2002) and, if they do, they may still report similar attitudes.

Box 7.4 Theft estimates in the workplace

The exact amount of theft that goes on in the workplace is hard to estimate given that it varies from one job to the next and, perhaps more fundamentally, that people are unlikely to report theft behaviours candidly. However, a variety of estimates have been made throughout the years. Hollinger and Clark (1983) estimated the non-trivial (i.e., deliberate, substantial and consequential) employee theft to be at least 5 per cent in most settings, though subsequent estimates have been much higher, particularly if they use alternative techniques to assess theft and other sensitive (illegal or counterproductive) behaviours (Dalton, Wimbush & Daily, 1994). Estimates have also varied across job sectors. For instance, Brooks and Arnold (1989) reckoned that for the retail industry theft rates are as high as 35 per cent, and Slora (1989) found that over 4/10 supermarket and over 6/10 fast-food employees confessed to theft (which suggests the real number is even higher). However, these estimates confounded both trivial and non-trivial theft (note that the former may include 'having at least once taken something from your work without paying'). However, studies controlling for methodological artefacts (such as underestimations due to self-reports) and distinguishing between more and less serious forms of theft are far from optimistic. Wimbush and Dalton (1997) estimated overall levels of theft to be as high as 60 per cent. Moreover, even non-trivial theft (e.g., more than $10 in cash or merchandise per a month) was estimated at 35 per cent.

Evidence for the validity of personality-based integrity inventories was reported quite emphatically by the largest meta-analytic study on personality constructs published to date (Ones, Viswesvaran & Schmidt, 1993). In it, Ones and colleagues reported a colossal total of 665 validity coefficients for over half a million subjects. The criteria used to examine the validity of integrity inventories included both self-reported and archival evidence of theft, attendance records and even job performance. The uncorrected average validity for integrity measures as predictors of job performance was .25 (corrected .41), whereas for counterproductive work behaviours (see also Section 6.11) it was .27 (corrected .39). This prompted experts to suggest that if integrity scales are added to GMA tests, the prediction of job performance is practically as good as it gets (Schmidt & Hunter, 1998).

Although critics have argued that most of the studies examined were published by the actual test-publishers (highlighting a potential conflict of interests, unless the publishers have integrity themselves), independent studies reported similar validities (Bernardin & Cooke, 1993). Yet it is true that when it comes to predicting actual theft, which is arguably the obvious criterion for validating any measure of integrity, the validities of integrity scales have been surprisingly low

(regardless of where the data came from) (Guastello & Rieke, 1991; Murphy, 2000).

7.18 Criticisms of integrity inventories

Like all psychological assessment tools, integrity inventories have been the target of criticisms, and perhaps even more so than other instruments because of the moral implications of assessing integrity. A recent review (Karren & Zacharias, 2007) has highlighted four major drawbacks of integrity scales, namely:

1. Construct confusion: it is unclear what integrity scales are actually assessing. If the underlying construct is integrity, why are these scales so weakly correlated with theft indicators, and why do they correlate quite systematically with job performance? From an applied perspective, integrity scales are especially needed to reduce or avoid theft (see Box 7.4), yet there are virtually no validity studies providing evidence for the idea that these scales are predictive of theft (Murphy & Dzieweczynski, 2005). Indeed, even in Ones' meta-analysis of integrity scales (see Section 7.17) different integrity scales were weakly intercorrelated, suggesting that different scales are assessing different constructs. Correlations of integrity with Conscientiousness, Agreeableness and Neuroticism (negatively) suggest that integrity is a higher-order or compound trait, but is 'integrity' the accurate name for it? On the basis of its Big Five correlates, one could relabel this trait something like 'Responsible pro-social stability' or 'Pro-active and stable cooperativeness'. It is apparent that the integrity–performance associations cannot be explained by Conscientiousness (Murphy & Lee, 1994), so why are integrity inventories predictive of performance? Finally, what do we know about the temporal and cross-situational stability of integrity, and is it more important in explaining dishonesty than situational factors are (Murphy, 1993)? Evidence suggests that integrity-related behaviours are more affected by situational than personal factors (Mumford, Connelly, Helton, Strange & Osburn, 2001). Indeed, one of the earliest reviews of personality (including integrity or 'character') measures concluded that different measures of honesty are very poorly intercorrelated (May & Hartshorne, 1926), such that people are honest or not depending on the situation and situationally determined motivation (see again Box 7.1).
2. Just like the polygraph (see Box 7.3), integrity scales may underestimate the number of false positive responses, misclassifying honest respondents as dishonest. As said (Section 7.13), 'faking good' about one's past misbehaviours may be picked up by lie scales – especially if one exaggerates. Yet being honest about them is unlikely to result in higher integrity. This problem brings us back all the way to Epimenides' paradox: if someone admits to being a liar, is that person really a liar? Moreover, it would only take a few fakers (who

are capable of not only spotting the correct response but also hiding their past misbehaviours or positive attitudes about immoral behaviours) to make honest respondents seem even more dishonest. It has been estimated that as many as four out of five respondents who end up being classified as dishonest are actually misclassified (Bernardin & Cooke, 1993).

3. In line with point 2, meta-analytic estimates suggest improvements of about 1 standard deviation on integrity 'scores' after test-takers have been instructed to fake good (Alliger & Dwight, 2000). This shows that, regardless of people's motivation to be honest, most individuals are able to identify the correct responses to these inventories. In theory, then, the only thing people would need to fake good is the motivation to do it (though of course some respondents may be particularly worried that they may be found out, especially in high-stake settings). Post hoc enquires to anonymous test-takers after completion of integrity scales indicate that up to 50 per cent of respondents admit having exaggerated or faked good, presenting themselves as more dependable, conscientious and reliable. Additional practice and coaching can cause substantial improvements on overt integrity tests and some improvements on personality-like inventories.
4. Employees and job candidates tend to dislike integrity inventories because of their intrusive nature (Ryan & Sackett, 1987) and the fact that they ask questions that bear little apparent relation to their view of competence. Indeed, many people regret taking such tests and dislike organisations for using them in selection. It has also been argued that integrity scales discriminate against people with certain political attitudes (Faust, 1996). For instance, agreeing with the statement 'It is important to give people a second chance' is seen as a sign of lack of integrity but it also taps into authoritarianism and right-wing personality, both negatively. It is therefore not surprising that people higher in Openness to Experience are more likely to 'fail' or score lower in integrity scales (Guastello & Rieke, 1991).

7.19 Emotional intelligence (EI)

In a broad sense, emotional intelligence (EI) refers to individual differences in the ability to identify and manage one's own and other people's emotions (Goleman, 1995; Salovey & Mayer, 1989). Thus, EI is essentially an ability related to emotional processes, notably the successful manipulation and accurate knowledge of emotional states. That said, just as integrity tests are not actual tests, EI is not really an intelligence; hence 'trait EI' is often used to emphasise the taxonomic position of this construct in the realm of personality rather than intelligence (Brody, 2004; Petrides & Furnham, 2001). But what is EI, and why is it not an actual intelligence? More importantly, how important is EI in personnel selection? Sections 7.20 to 7.23 attempt to answer these questions.

7.20 What is EI?

As said, EI refers to individual differences in emotional identification and management, hence it is generally defined as the ability to identify and manage one's own and others' emotions. However, there has been extensive debate in the past ten years or so as to whether the EI construct fulfils the necessary requirements to be considered an actual intelligence (for a state-of-the-art summary of this debate and a compelling case against the inclusion of EI in the realm of human intelligence the reader is strongly encouraged to consult Brody, 2004). Note that the use of the term 'ability' is neither consistent with how EI tends to be assessed nor sufficient to grant EI a place in the realm of human abilities. Unless, say, we consider any self-report of competence an indicator of intelligence. For example, some people may be better than others at making other people cry, and we may assess these individual differences by asking respondents to agree or disagree with statements such as 'I find it very easy to make others cry' or 'Upsetting others is easy.' But would these questions be measuring an *ability*, let alone intelligence?

Whilst there is no absolute answer to these questions, it is clear that if one conceptualises EI as an intelligence it would be incongruent not to do the same with a wide range of other non-cognitive constructs. In fact, most personality traits would probably have to be rebranded as 'intelligence'. For example, Emotional Stability (low Neuroticism) could be defined as the ability to remain calm under difficult circumstances or the ability to work well under pressure, Extraversion could be defined as the ability to socialise or interact with people even if you don't know them, and Conscientiousness could be defined as the ability to resist one's temptations and trade off instant gratification for relevant long-term goals. Doing this, however, would be incongruent with a well-established tradition in psychology to refer to maximal performance tests as 'ability tests' (see Section 7.1) and typical performance scales as 'personality inventories'. In that sense, there are two radical elements in the concept of EI (though some approaches, such as trait EI, are an exception to this; see Section 7.21), namely (a) it defies the notion of 'intelligence' as being essentially cognitive, and (b) it defies the notion of ability tests as being essentially maximal performance measures. In fact, a third important issue is that ability tests are predictive of individual differences in learning, educational and occupational achievement, where occupational achievement is particularly relevant to this book. Can the same be said about EI scales? Do they predict job and training performance?

7.21 EI: the personality construct

Conceptualisations of EI as a personality construct are congruent with the traditional distinction in psychology of referring to self-reported scales of

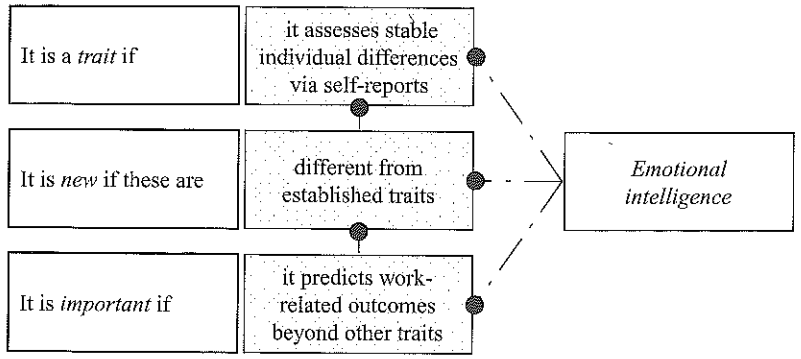


Figure 7.23 *Validating emotional intelligence as a personality construct: three 'ifs'*

typical performance as personality traits and, accordingly, seek to validate scales of EI as personality measures. First, this requires demonstration of the reliability of EI self-reports, that is, showing that these scales have sufficient test–retest reliability to indicate little intra-individual variability and a 'trait-like' disposition. Second, it is also essential to demonstrate that the EI construct is different from established personality traits, otherwise EI would simply be a new name for an already known trait or set of traits (Chamorro-Premuzic & Furnham, 2006a). Third, and perhaps most importantly from an applied point of view, it is important to show that any association between EI measures and relevant outcomes (e.g., job performance, job satisfaction and training success) cannot be explained by established personality traits, such as the Big Five. Thus EI should demonstrate *incremental validity* over and above other traits in the prediction of work-related outcomes.

7.22 Reliability of EI

There is no question that the constellation of behavioural tendencies defined by the EI construct refers to individual differences in stable behavioural patterns, such that one may safely conclude that what EI inventories assess is personality. Moreover, they found that trait EI is significantly correlated with established personality traits. These results provide compelling evidence for the interpretation of EI as a personality trait that is a 'compound' or higher-order factor of individual differences in Emotional Stability, Extraversion, Agreeableness and Conscientiousness (see also Section 7.23). Petrides and Furnham's (2001) original interpretation of 'emotional self-efficacy' emphasises the fact that, from a conceptual point of view, people who are less Neurotic, more Extraverted, Agreeable and Conscientious *rate* their abilities to handle emotionally challenging situations and their abilities to perceive and manage emotions higher than their more neurotic, introverted, disagreeable and less conscientious counterparts. In the tradition of self-expectancy constructs, such as self-efficacy (see

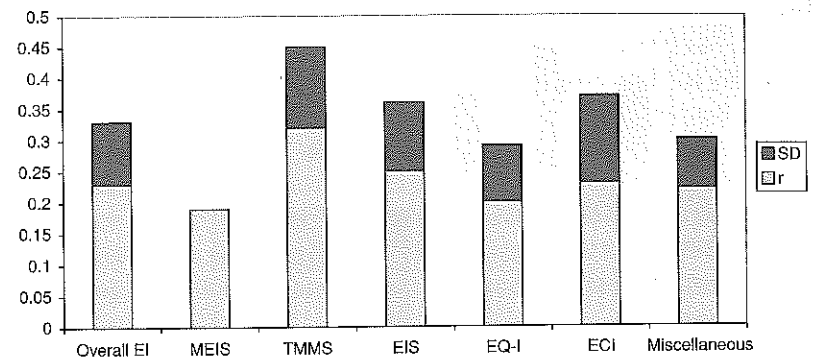


Figure 7.24 *Meta-analytic validities for different EI scales (corrected correlations and their SDs)*

Section 7.24), people's perceptions of their own abilities (emotional or not) are known to have important behavioural consequences regardless of their accuracy. Thus the question for trait EI is whether it can have self-fulfilling effects on performance related to work outcomes; so, just how valid are self-reports of EI in predicting occupational success? (See Section 7.23).

7.23 Validity of EI

The importance of EI in the context of work was demonstrated by a meta-analytic study (Van Rooy & Viswesvaran, 2004). The authors examined data from 59 independent studies (9,522 participants) and reported an operational validity of .23 for measures of EI as predictors of various job outcomes. Associations between EI and the criteria examined depended on what instruments were used (Figure 7.24) and what dimensions or aspects of EI were assessed (Figure 7.25). As seen in Figure 7.24, validities were slightly higher for the trait meta-mood scale (TMMS) (Salovey, Mayer, Goldman, Turvey & Palfai, 1995), followed by the emotional intelligence scale (EIS), emotional competence inventory (ECI) (Sala, 2002), miscellaneous scales, the emotional quotient inventory (EQ-I) (Bar-On, 1997) and finally the multifactor emotional intelligence scale (MEIS) (Mayer & Salovey, 1997).

Figure 7.25 plots the validities of different 'facets' or aspects of EI as estimated by Van Rooy and Viswesvaran's meta-analysis. The four bars on the left represent subscales of the MEIS (based on Salovey and Mayer's framework) and show that assimilation (7 samples and 770 participants) had the highest and perception (21 samples and 3,484 participants) the lowest validities, and that understanding (10 samples and 1,525 participants) and management (18 samples and 2,961 participants) showed the highest variability (as indicated by the SD of the meta-analytic correlations). The five bars on the right of the chart are based on the EQ-i (Bar-On's) subgroupings and indicate that validities were generally lower

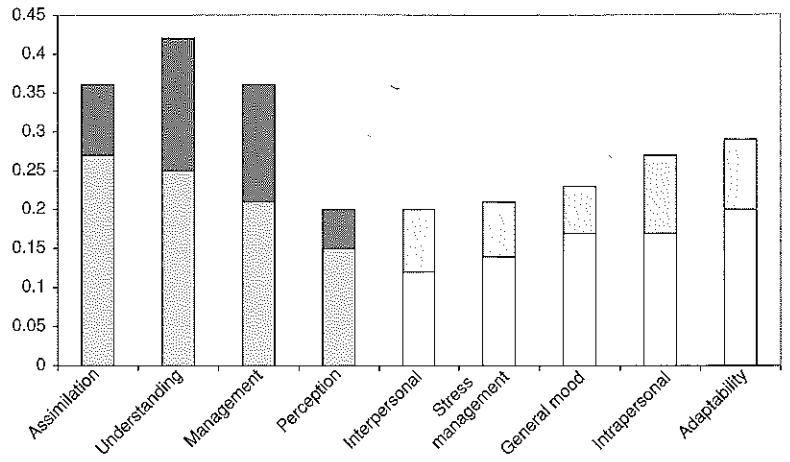


Figure 7.25 *Aspects of EI that predict work outcomes*

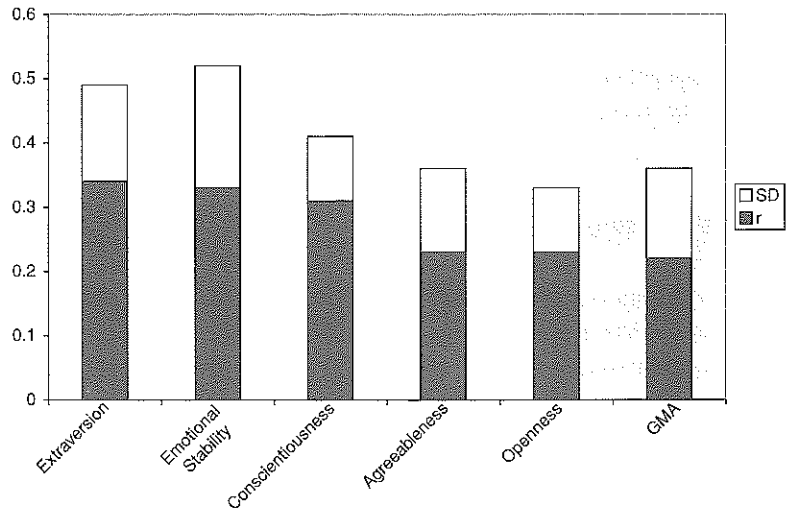


Figure 7.26 *Meta-analytic correlations (and their SDs) of EI with the Big Five and GMA*

and more homogeneous than for the MEIS subscales. The lowest validity was found for the interpersonal facet (22 studies and 4,684 participants), followed by stress management (9 samples and 2,687 participants), followed by general mood (9 samples and 2,687 participants) and intrapersonal (20 studies and 4,548 participants); the highest validity was found for adaptability (20 samples and 4,524 participants).

Van Rooy and Viswesvaran also reported meta-analytic correlations between EI, GMA and the Big Five personality traits. As shown in Figure 7.26, EI correlated at .34 with Extraversion (19 studies and 4,158 participants), 33 with Emotional Stability (23 studies and 4,213 participants), .23 with Openness and Agreeableness (14 studies and 3,306 participants), and .22 with GMA (19 studies

and 4,158 subjects). All correlations correct for unreliability in the criterion and observed mean in the predictors.

Moreover, the authors used previous validities for GMA and the Big Five to estimate the incremental validity of EI, and found that EI explained unique and additional variance over and above personality traits, but not GMA. Other studies have also shown that EI inventories are valid predictors of peer nominations of influence (Byrne, Smither, Reilly & Dominick, 2007), though a recent study showed that EI has no incremental validity over GMA and Conscientiousness in predicting educational attainments, social status or income (Amelang & Steinmayr, 2006).

7.24 Self-efficacy

Another personality construct that has long been associated with work performance is self-efficacy, defined broadly as an individual's beliefs about his or her capabilities to affect events and outcomes and produce desired levels of performance (Bandura, 1999). The core idea underlying the notion of self-efficacy is that competency-related self-beliefs have self-fulfilling prophecies and affect performance irrespective of actual abilities. To paraphrase Henry Ford, 'Whether you believe you can do a thing or not, you are right.'

Although self-efficacy has been primarily examined in the context of clinical and social-psychological settings, there is a wealth of published research looking at self-efficacy in work settings too. Judge, Jackson, Shaw, Scott and Rich (2007) note that: 'In industrial–organisational (I-O) psychology, self-efficacy has been remarkably popular as well. In the past 25 years, more than 800 articles on self-efficacy have been published in organisational journals. Virtually every area in organisational research has utilized self-efficacy' (p. 107). Thus Judge *et al.* provided a quantitative estimate of the unique contribution of self-efficacy to work-related outcomes whilst taking into consideration the Big Five traits, GMA and previous job experience. Results showed that overall the contribution of self-efficacy to explaining occupational outcomes was relatively small in comparison to that of other personality traits and GMA. However, there were some exceptions, notably low complexity jobs (this is consistent with Kanfer & Ackerman, 1989) and task rather than job performance, where self-efficacy did provide some important additional information to the prediction of work outcomes (Judge *et al.*, 2007). However, in all the effect of self-efficacy was largely accounted for by personality and GMA.

7.25 Core self-evaluations

The idea that self-beliefs affect performance independently of actual competencies has not only been represented by the notion of self-efficacy but is

at the heart of a wide range of self-centred constructs, such as locus of control, self-esteem, self-concept and even Neuroticism. In an attempt to overcome the multiplicity of labels and overlapping constructs referring to individual differences in self-perceptions (e.g., of concept, ability, efficacy, etc.), Judge and colleagues recently conceptualised a higher order construct of 'Core Self-Evaluations' (CSE) (Judge, Erez, Bono & Thoresen, 2003; Judge, Locke, Durham & Kluger, 1998). This work was inspired by the authors' finding (see Section 7.11) of the strong dispositional determinants of job satisfaction. But the key importance of CSE is that it explains the common variance found among a variety of self-centred constructs, such as locus of control, self-efficacy, self-esteem, self-concept and also Neuroticism. Thus high CSE are indicative of positive self-regard in general, and specifically higher self-efficacy, internal rather than external locus of control or attributional style, Emotional Stability rather than Neuroticism, and higher self-concept. Evidence indicated that people with higher CSE are also more likely to hold self-concordant goals and be intrinsically rather than extrinsically motivated, all of which is reflected in higher job satisfaction (Judge, Bono, Erez & Locke, 2005).

7.26 Moving beyond traits: the person–environment fit

Though not new (Pervin & Rubin, 1967), interactive approaches to personality have emphasised the importance of assessing not only traits but also situational factors, specifically the level of congruence between them. Thus such theories posit that occupational outcomes, such as job satisfaction, stress and indeed productivity, will depend not only on the personal characteristics (e.g., Conscientiousness, Integrity or Extraversion) but on whether these *fit* in the context of the organisation (Pervin, 1989). Accordingly, whether specific personality traits affect occupational outcomes and how will depend on whether the environment facilitates or inhibits their effects. For example, Extraversion may be predictive of higher performance in sales jobs only if these environments are sufficiently 'extraverted'. What matters here is the individuals' perception of the environment as well as his or her assessment of how close that environment is to his or her 'ideal' setting. In that sense, three things are needed to assess the level of person–environment fit, namely:

(a) The individual's personality
(b) The environment's 'personality'
(c) The level of congruency between a and b

Recent evidence suggests that the person–environment fit predicts training success in occupational settings. Specifically, a study found that if employees perceived that their training supervisors supported and promoted their creative initiative and respected their preferred workload, they performed better than employees

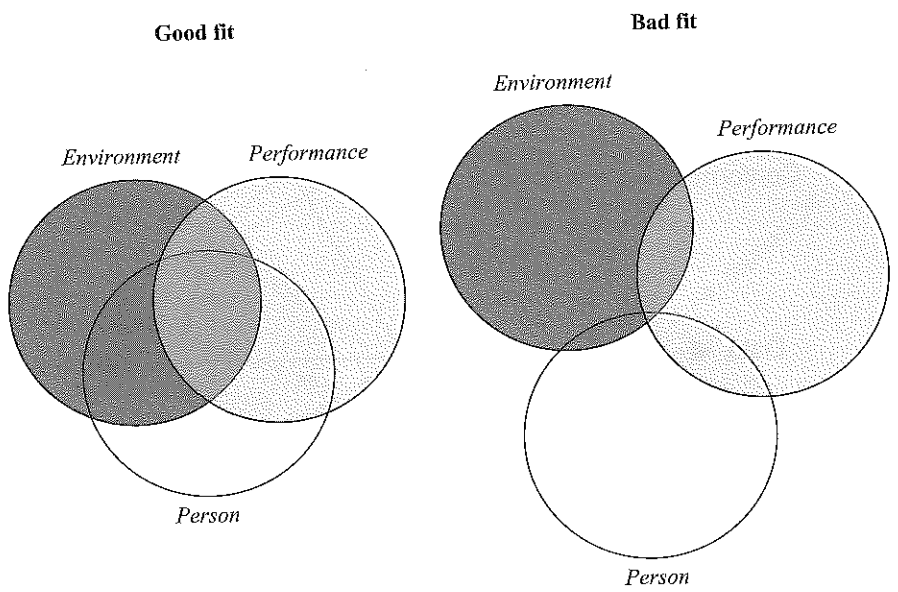


Figure 7.27 *Performance as a function of the person–environment fit*

who did not feel supported in these areas by their supervisors (Awoniyi, Griego & Morgan, 2002).

Although assessment of environmental factors has been largely overshadowed by the assessment of personality traits, the vocational interests – or simply 'interests' – have embodied a longstanding attempt to bridge the gap between personal and environmental attributes. Within this field, major taxonomies have emerged such as Holland's and Prediger's interests models, which are often conceptualised in terms of personality traits (Holland, 1999). Holland (1959, 1997) classified individuals' vocational preferences and indeed the level of prestige, income and skills associated with these vocations, using a typology that could be simultaneously applied to people and jobs, hence emphasising the importance of 'fitting' the right person to the right job. As seen in Figure 7.28, Holland's typology classifies interests and careers with an hexagon describing Realistic (R), Investigative (I), Artistic (A), Social (S), Entrepreneurial (E) and Conventional (C) jobs/interests, with R-S representing the Things–People dimensions, C-A representing the Conformity (non-Conformity) dimension, and the Data-Ideas and Sociability dimensions at the intersection of C-E/I-A and E-S/R-I, respectively. People's interests have been shown to be at least as reliable as standard personality traits, if not more. For example, Strong reported a twenty-two-year (eighteen years after university) test–retest reliability of .75 for his Interest Inventory (Strong, 1955).

A recent study (Armstrong *et al.*, 2008) reported important data highlighting the interface between personality (e.g., Big Five) and interests (Holland's RIASEC), as well as examining sex differences in these constructs. Results

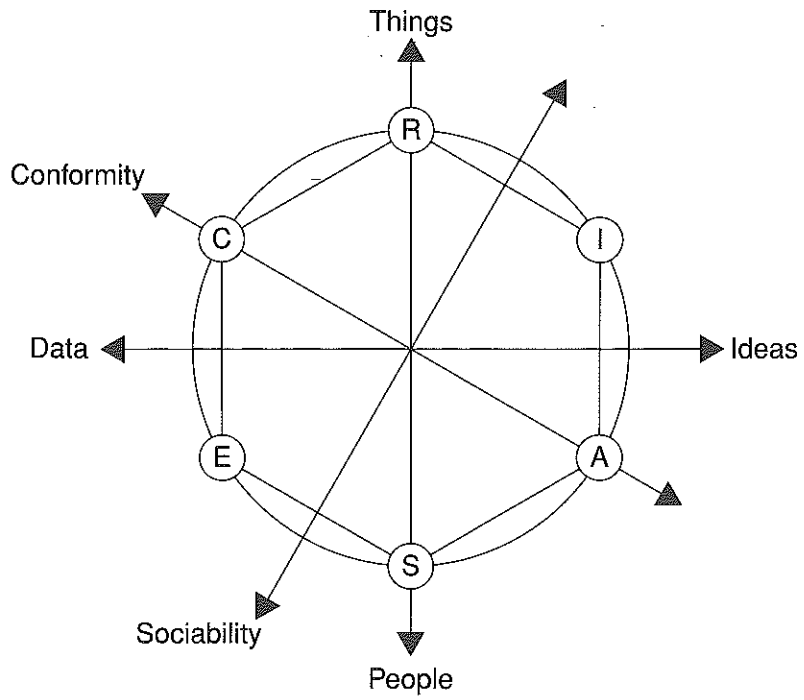


Figure 7.28 *Holland's RIASEC*

indicated that men showed a preference for 'Things' whereas women showed a preference for 'People', which is in line with previous studies (Tracey & Rounds, 1992). Indeed, Lubinski argued that this sex difference is the 'largest of all sex differences on major psychological dimensions' (2000, p. 421). But the most interesting results of Armstrong *et al.*'s investigation concern the associations among different individual difference constructs. As shown in Figure 7.29, different personality dimensions could be organised on the basis of the RIASEC model in a way that is coherent with the meaning of the constructs examined.

7.27 Summary and conclusions

Throughout this chapter we have examined the notion of personality traits and how self-report-based inventories designed to assess them can be used to predict important occupational outcomes. As we have seen, personality traits differ mainly from cognitive ability constructs in that the former are assessed via self-report inventories and refer to individual differences in typical behaviours, whilst the latter (see Chapter 6) measure maximal performance and are measured via timed tests of performance (with objective rather than subjective items). Although the predictive power of personality scales at work is clearly lower and more job-specific than that of ability tests, if the right traits are assessed in the right

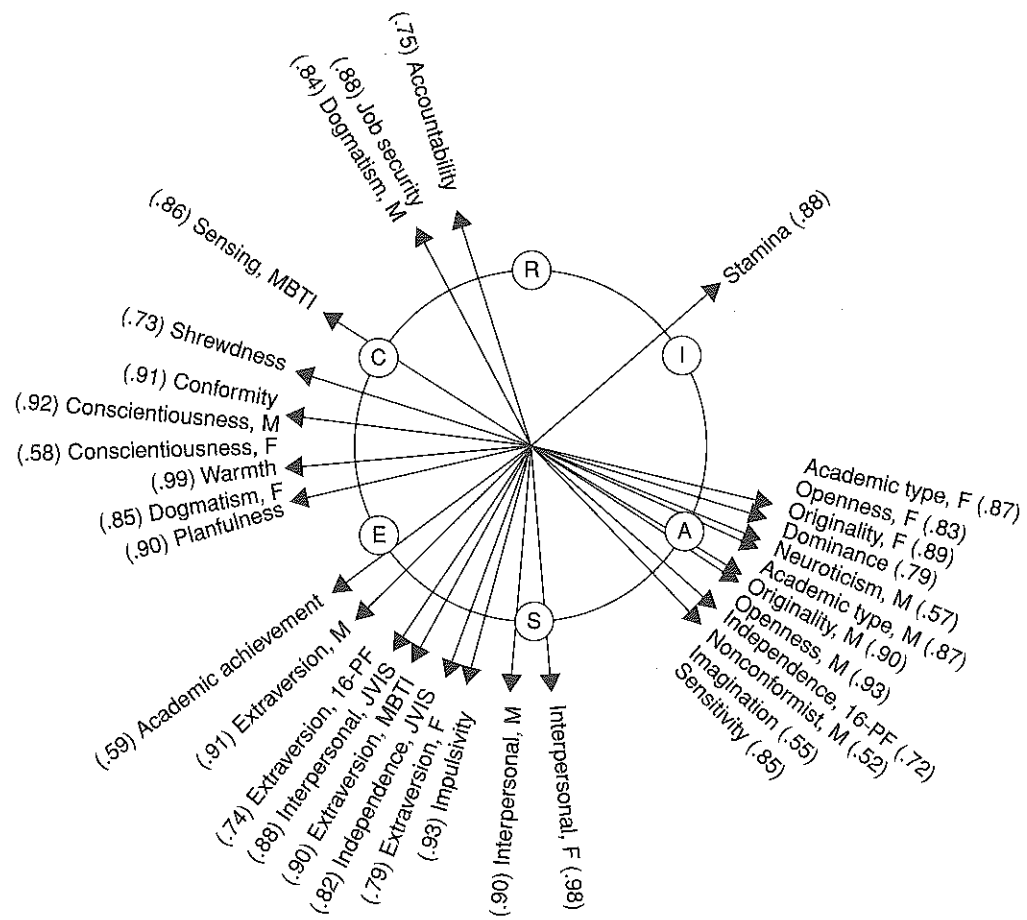


Figure 7.29 *Individual characteristics integrated into a two-dimensional RIASEC interest circumplex (reproduced with permission from Armstrong* et al., *2008)*

context, personality inventories add important information to the prediction of future occupational success. Moreover, Conscientiousness, one of the main factors of personality, has general and positive effects on work-related criteria that are independent of GMA as well as preventive and protective health- and real-life-related effects.

Although the face validity of self-reports is threatened by the fact that in high-stake settings people can (and are indeed motivated to) 'fake good' or lie, there is mixed evidence and almost never-ending debate as to the real practical problems that faking poses. What is clear is that we have no 'objective' (ability-like) alternatives for measuring individual differences in personality, and that self-reports still provide incremental validity in the prediction of work outcomes over and above GMA.

Several steps forward have been taken in personality-related personnel selection since the early 1990s. First, meta-analytic estimates have provided reliable and robust quantitative evaluations of the predictive validity and utility of widely used traits. Second, personality researchers have at last agreed on the major personality traits that should be considered to provide a comprehensive picture of individual differences in non-clinical samples, with the Big Five emerging as the dominant taxonomy. Although the emergence of the Big Five should not undermine other models, it is inarguably useful as a universal 'exchange currency' to compare results from different studies (which, for instance, in the case of meta-analytic validations, is of capital importance). Third, 'compound' traits, such as Integrity, EI and self-centred constructs (e.g., self-efficacy or CSE) are also useful explaining non-ability-related variance in occupational outcomes. Last, but not least, recent years have witnessed an exciting attempt to integrate core personality constructs with their 'cousin' constructs of interests and even abilities (Ackerman, 1997; Armstrong *et al.*, 2008; Chamorro-Premuzic & Furnham, 2006). These conceptual efforts are important because of their potential to capitalise on the synergistic (*mediational* and *moderational*) links among the major sources of interindividual differences, in the workplace and beyond.

Notes

1 Throughout this chapter we will use 'subjective' to encompass both self- and other-reports of traits or behaviours, purely in order to distinguish between this form of assessment and 'objective' tests (which have objectively determined correct answers).
2 See www.personality-project.org/perproj/theory/big5.table.html for details.

8 Creativity

8.1 Introduction

Some employees are more creative than others: they are more likely to come up with original thoughts and novel solutions and stand out in organisations for their innovative thinking; they seem to prefer innovation to imitation and even enjoy defying the crowd. In this chapter, we discuss the key psychological factors underlying creativity and what personnel selection can do to select creative people.

Although the topic of creativity has a longstanding history in psychology (dating back to the very beginnings of intelligence testing more than one hundred years ago), creativity researchers have repeatedly complained about the fact that insufficient attention is given to the field (Guilford, 1950; Sternberg & Lubart, 1996). Indeed, despite growing economic interests and being associated with a wide range of concepts, such as intelligence (see Chapter 6), personality (see Chapter 7), leadership (see Chapter 9), imagination, motivation, social influence, intuition and talent (see Chapter 10) (Runco, 2004), creativity continues to be neglected from selection-related research.

In 1950, Guilford highlighted the importance of increasing creativity research after noting that only 186 of the 121,000 psychological studies in databases had dealt with creativity. In line with this finding, Figure 8.1 shows that the number of articles including 'creativity', 'creative' or 'originality' as keywords in the main applied journals (*JAP* = *Journal of Applied Psychology*; *JCP* = *Journal of Counselling Psychology*; *JOB* = *Journal of Organizational Behavior*; *AMR* = *Academy of Management Review*; *PP* = *Personnel Psychology*; *JOOP* = *Journal of Organizational and Occupational Psychology*; and *IJSA* = *International Journal of Selection and Assessment*) is very low, even for *JAP*, as the total number of articles published by the journal between 1917 and 2008 was 8,932. Needless to say, a great deal of creativity research has been published in other, creativity-specialised journals, such as the *Journal of Creative Behaviour* and *Creativity Research Journal*; however, these publication outlets are rarely concerned with personnel selection and tend to examine student samples.

The generalised lack of applied research on creativity is in stark contrast with the consensus, particularly in industrialised or developed nations, on the importance of investing in creative employees. Thus Porter (1990, p. 73) noted that 'national prosperity is created, not inherited' and Amabile (1990) saw individual