

## The General Factor of Personality and Evaluation

ERIK PETTERSSON\*, ERIC TURKHEIMER, ERIN E. HORN and ANDREW R. MENATTI

Department of Psychology, University of Virginia, Charlottesville, VA 22904-4400, USA

**Abstract:** According to the proposal of the general factor of personality (GFP), socially desirable personality traits have been selected for throughout evolution because they increase fitness. *However, it remains unknown whether people high on this factor actually behave in socially desirable ways or whether they simply endorse traits of positive valence.* We separated these two sources of variance by having 619 participants respond to 120 personality adjectives organised into 30 quadruples balanced for content and valence (e.g. unambitious, easy-going, driven and workaholic tapped the trait achievement-striving). An exploratory six-factor solution fit well, and the factors resembled the Big Five. We subsequently extracted a higher-order factor from this solution, which appeared similar to the GFP. *A Schmid–Leiman transformation of the higher-order factor, however, revealed that it clustered items of similar valence but opposite content (e.g. at the negative pole, unambitious and workaholic), rendering it an implausible description of evolved adaptive behaviour. Isolating this evaluative factor using exploratory structural equation modelling generated factors consisting of items of similar descriptive content but different valence (e.g. driven and workaholic), and the correlations among these factors were of small magnitude, indicating that the putative GFP capitalises primarily on evaluative rather than descriptive variance.* Implications are discussed. Copyright © 2011 John Wiley & Sons, Ltd.

**Key words:** general factor of personality; Big One; evaluation; valence; social desirability

Whether quantified as the first principal component (PC), an unrotated factor accounting for the most variance in an exploratory factor analysis (EFA) or as the top level in a higher-order factor model, the most important dimension in self-reported and other-reported personality responses contrasts positively evaluated items against negatively evaluated items. This is not a new observation. A perusal of the history of scientific study of the structure of personality shows that the issue of how positively versus negatively evaluated items should be incorporated into the structure of personality responses has come up again and again but has never been fully resolved.

The evaluative dimension may be operationalised as the perceived valence attached to a given item or descriptor, irrespective of its content, that is, what behaviour, thought or feeling that it is purported to measure. In the past, this general dimension has been conceptualised as *widely different constructs, including anxiety (Welsh, 1956), social desirability (Edwards, 1957) and resiliency (Block, 1965).* More recently, this dimension has resurfaced as the general factor of personality (GFP; Musek, 2007; Rushton, Bons, & Hur, 2008), which, in contrast to previous theories, attributes evolutionary importance to traits of positive valence. According to the GFP,<sup>1</sup> individuals who were agreeable, conscientious,

emotionally stable, extraverted and open to experience succeeded to survive and reproduce better than those who were disagreeable, lazy, neurotic, introverted and closed to experience. Three lines of evidence support this theory. First, many studies have reported a general factor whose defining traits seem to be of similar valence (Musek; Rushton & Irwing, 2008; Rushton & Irwing, 2009a, 2009b, 2009c; Rushton, Bons, Ando, Hur, Irwing, Vernon, et al., 2009).

Second, Rushton et al. (2008) postulated that the GFP is similar to intelligence in that higher scores are always beneficial because they increase fitness. From a heritability point of view, such constructs are governed under unidirectional selection, which manifests when identical twins are more than twice as similar compared with fraternal twins, that is, when combinations of genes are more important than their cumulative sum. This type of heritability involves non-additive or dominance effects, and evidence that the GFP displays this type of heritability has emerged in three different twin samples (Figueredo & Rushton, 2009; Rushton et al., 2009). Although one might expect that variation in traits governed under dominance effects would eventually disappear, it may persist for at least two reasons. First, mutation load may operate as a countering force, balancing out the effect of the relative reproductive success of individuals high in GFP (Penke, Denissen, & Miller, 2007). Second, Rushton et al. (2008) argued that just as intelligence is negatively affected by inbreeding, so is the GFP, citing evidence that inbred families demonstrate lower scores on extraversion, openness and sensation seeking.

\*Correspondence to: Erik Pettersson, University of Virginia, Department of Psychology, PO Box 400400, Charlottesville, VA 22904-4400, USA. E-mail: etp5v@virginia.edu

<sup>1</sup>In many studies, the GFP is operationalised as the primary factor of a given inventory, which, as noted by an anonymous reviewer, is somewhat arbitrary. More recently, researchers have begun to develop questionnaires in order to measure this dimension in a more consistent fashion (Amigo, Caselles, & Mico, ).

Third, some behavioural predictions based on the GFP have been supported. According to theory, the GFP should covary with constructs that may pertain to reproductive success. In one study, a general measure of intelligence correlated positively with the GFP at a medium effect size (Schermer & Vernon, 2010). In addition, self-reported GFP scores correlated positively at small to medium effect sizes with several work performance indices as evaluated by supervisors (van der Linden, te Nijenhuis, & Bakker, 2010). The GFP has also been found to relate to high school likability and popularity in a Dutch sample (van der Linden, Scholte, Cillessen, te Nijenhuis, & Segers, 2010).

Thus, to summarise, there are some supports for the GFP theory in that scales of similar valence tend to cluster, dominance effects are present in heritability studies, and the GFP relates to substantive outcomes presumed to impact general reproductive success such as intelligence and popularity. However, the GFP has also failed to receive support in some of these domains. With regards to genetic dominance, Veselka, Schermer, Petrides, and Vernon (2009) failed to find support for genetic dominance in two twin samples, which they attributed to lack of power. Remedying this problem, Loehlin and Martin (2011) analysed two Australian twin samples each consisting of several thousands of pairs. They demonstrated that although the GFP evidenced some dominance effects, so did the scales from which the GFP had been statistically removed, suggesting that dominance was not unique to the GFP *per se*. With regards to the structure of the GFP and its predictive abilities, de Vries (2011a) re-analysed previously published data used to advocate the presence of the GFP but found that models without the GFP fit the data equally well (cf. Ashton, Lee, Goldberg, & de Vries, 2009), and that the GFP did not add any predictive power to high school popularity and likeability once extraversion was controlled for first (de Vries, 2011b).

### Evaluation and the general factor of personality

Aside from these concerns, the relation between the GFP and social desirability remains controversial. As Edwards (1957) noted, loadings on the first PC of self-reported personality and psychopathology inventories correlate almost perfectly with independent ratings of the social desirability of the items. In other words, people tend to endorse items that sound positive and deny items that sound negative, suggesting that part of the covariance among items and scales may be attributable to shared valence, rather than descriptive content. Bäckström (2007) addressed this by demonstrating that once evaluation was statistically controlled, the magnitude of the correlations among the Big Five factors decreased substantially. Furthermore, after a set of Big Five items was rewritten in a less evaluative fashion (for example, the item 'Feel little concern for others' was changed to 'Believe it is better if everyone cares for himself or herself'), the covariance among items decreased substantially (Bäckström, Björklund, & Larsson, 2009). In a confirmatory factor analysis (CFA), the original Big Five items loaded strongly on a general factor (mean loading=0.56), whereas the rewritten,

non-evaluative items loaded only weakly on a general factor (mean loading=0.09).

However, the meaning of these studies for the theory of the GFP remains somewhat unclear. It may be that the part of extraversion that pertains to positive valence correlates with actual behaviour and that isolating this variance artefactually decreases extraversion's relation with other traits of positive valence such as conscientiousness and agreeableness (cf. Heilbrun, 1964). Another way to address whether the covariance among traits of similar valence represents some kind of artefact or bias is to have several raters describe a single target. To the extent that raters agree, that may be considered a substantive variance because raters may not be expected to share idiosyncratic biases. In contrast, variance not shared among raters may be more likely to represent biases or error.

Taking this approach, Anusic, Schimmack, Pinkus, and Lockwood (2009) observed that the correlation between two higher-order factors, alpha (which consisted of loadings on emotional stability, agreeableness and conscientiousness) and beta (which consisted of loadings on extraversion and openness to experience), became non-significant when a general halo factor was controlled, replicating the works of Bäckström (2007) and Bäckström et al. (2009). Moreover, the halo factor correlated close to zero among observers. DeYoung (2006) similarly demonstrated that although the correlation between alpha and beta was significant at an individual level, once specific rater effects was controlled, the correlation became non-significant. Moreover, McCrae, Yamagata, et al. (2008), in analyses of twins for whom there were both self-report and other-report data, found evidence that covariation among traits represented primarily idiosyncratic variance unique to only one twin within a pair or unique to one rater but not shared by independent observers. Thus, the most parsimonious models in their analyses indicated that trait covariance was primarily due to target or rater idiosyncrasies that may not represent substantive variation.

However, it should be noted that idiosyncratic variance unique to a single twin or observer is not necessarily bias or error. As has been observed previously, 'any particular informant may accurately detect some aspect of the target's personality that other informants have overlooked,' or 'Individuals may know things about themselves, through introspection, that others do not' (DeYoung, 2006, pp. 1147). Likewise, McCrae et al. (2008) noted, 'Although personality psychologists usually interpret agreement as evidence of accuracy and disagreement as evidence of method bias, neither of these is necessarily the case: Agreement may be false consensus, and disagreement may reflect unique knowledge' (pp. 452). Rater-specific trait covariance, on the other hand, appears likely to represent nuisance variance. In this manuscript, we will isolate social desirability using a different method, with the goal of exploring the basis of the GFP. This method consists of developing item pairs balanced for evaluative and descriptive content.

### Balanced item sets

Originally introduced by Peabody (1967) to separate evaluative from descriptive variance, this method consists of

developing quadruples of items tapping a single trait matched for both valence and description. For example, the items *stingy*, *thrifty*, *generous* and *extravagant* are claimed to measure an underlying trait akin to generosity. This set allows participants to endorse items based on valence as well as description. Someone who is at the lower end of this continuum but has a tendency to, for whatever reason, endorse items of positive valence might subscribe to being *thrifty* but not *stingy*. A person at the same end of the continuum but with a tendency to endorse negatively valenced items might endorse *stingy* but not *thrifty*. Hence, the advantage of these balanced item sets is that they allow for separating valence from description, which has merit in investigations of the evaluation-loaded GFP.

One investigation took advantage of Peabody's (1967) item sets to tease apart whether the first PC of self-report responses measured substantive variation or social desirability (Edwards, 1969). If this dimension clustered items by description, then items such as *thrifty* and *stingy* should load together. If, on the other hand, this dimension clustered items by valence, then items such as *stingy* and *extravagant* should load together. It turned out that the latter was true, that is, that the first PC consisted of items of similar valence but opposite descriptive content. For example, someone low on this dimension self-described himself/herself as *stingy* and *extravagant*, *inhibited* and *impulsive*, *grim* and *frivolous*, *tense* and *lethargic*, *self-disparaging* and *conceited*, and *secretive* and *indiscreet*. Thus, valence, rather than content, was the most important way in which people responded to Peabody's items (as measured by variance accounted for).

Based on these studies, we submit that there are two broad classes of reasons why a respondent might tend to give predominantly negative-valenced or positive-valenced responses to personality items. The first is that the items might in fact be valid indicators of the respondent's behaviour. It might be the case, as Rushton et al. (2008) have suggested, that there is a broad evolutionary tendency to favour adaptive over non-adaptive personality patterns, so all adaptive traits are positively correlated with each other. The second class of reasons involves what might be termed evaluative response bias; rather than corresponding to the actual behaviour of a respondent, the tendency to endorse positively or negatively evaluated items might reflect a general self-evaluation that is otherwise independent of the descriptive content of the items. There are many familiar examples of similar biases, including acquiescence bias (a tendency to respond affirmatively to items), faking either good or bad, or for that matter responding randomly. The point is not that any of these biases occur by chance. They are no doubt adopted by certain respondents for particular reasons (e.g. the need to belong, Crowne & Marlowe, 1964) and will therefore be correlated with other individual differences that may be of considerable interest. What they are not, however, are valid indicators of the descriptive content of the items themselves. If a respondent endorsed the choice in the right column for all items, it might say something interesting about him or her, but the experimenter or clinician would not interpret the content of the endorsed items as a literal description of the respondent's personality.

We propose that the actual role of evaluative variation in self-reported scores on the GFP has never been definitively established. Is the evaluation dimension a fundamental axis of human behaviour with evolutionary underpinnings, a methodological response bias or a combination of the two? To resolve this question, we utilise balanced item sets inspired by Peabody (1967) to examine the role of evaluation in the GFP. Furthermore, in contrast to Edwards (1969), we investigate areas beyond evaluation using exploratory structural equation modelling (ESEM; Asparouhov & Muthén, 2009).

## METHOD

### Participants

The first set of participants consisted of 32 (18 women) individuals enrolled in introductory psychology classes at a state university on the East coast of the USA. Their average age was 18.66 ( $SD=1.07$ ; range=16–21 years). Of the first set of 32 participants, 24 identified themselves as White, 2 as Asian, 1 as African–American and the rest as other.

The second set of participants consisted of 619 individuals (390 women) enrolled in introductory psychology classes at the same university as the first set of participants. The mean age for this sample was 18.74 ( $SD=1.22$ ; range=17–36 years). Of the second set of 619 participants, 387 identified themselves as White, 104 as Asian, 45 as African–American and the rest as other.

### Materials and procedure

We constructed a set of 120 adjectives organised into 30 quadruplets corresponding to the 30 facets of the NEO PI-R (Costa & McCrae, 1992). We selected the NEO PI-R facets because they cover most aspects of the personality domain rather well. The adjective set is presented in Table 1. Following Peabody (1967), for each facet, we selected four adjectives, two hypothesised to have positive evaluation and two with negative evaluation, representing opposite poles of the facet. For example, for the Activity facet, we included adjectives *sluggish*, *placid*, *active* and *manic*.

The first set of 32 participants rated how desirable it would be for an individual to exhibit a given trait as indicated by the adjectives, ranging from 1 (not at all desirable) to 5 (very desirable). The second set of 619 participants provided self-report responses to the 120 items. The self-report items were answered on a 5-point Likert scale, ranging from 1 (not at all like me) to 5 (very much like me).

## RESULTS

### Evaluation ratings

We used the intraclass correlation coefficient function in the 'psych' package (Revelle, 2009) in the statistical package R (R Development Core Team, 2010) to compute a reliability coefficient for the mean evaluation rating of the 32 raters who rated each item. The average reliability of item ratings equaled 0.99, indicating a very high degree of convergence

Table 1. Items, percent evaluative variance, and Schmid-Leiman *g* loadings

Facet	Negatively evaluated			Positively evaluated			Positively evaluated			Negatively evaluated		
	low trait	% Eval	SL	low trait	% Eval	SL	high trait	% Eval	SL	high trait	% Eval	SL
(E) Warmth	Cold	53	-0.41	Reserved	7	0.09	Warm	22	0.44	Gushy	17	-0.06
(E) Gregariousness	Withdrawn	68	-0.19	Private	5	0.01	Sociable	37	0.15	Clingy	57	-0.14
(E) Assertiveness	Submissive	38	0.00	Modest	10	0.40	Assertive	11	-0.10	Overbearing	46	-0.21
(E) Activity	Sluggish	76	-0.35	Placid	0	0.09	Active	27	0.20	Manic	64	-0.37
(E) Excitement-seeking	Frightened	62	-0.14	Cautious	1	0.24	Bold	9	-0.11	Rash	52	-0.38
(E) Positive Emotion	Grim	82	-0.38	Serious	0	0.05	Cheerful	9	0.37	Frivolous	70	-0.25
(A) Trust	Suspicious	60	-0.27	Sceptical	29	-0.23	Trusting	42	0.30	Gullible	22	-0.01
(A) Straightforwardness	Secretive	37	-0.11	Discreet	0	0.15	Frank	1	-0.14	Indiscreet	15	-0.21
(A) Altruism	Selfish	77	-0.45	Self-reliant	14	0.06	Altruistic	28	0.25	Self-effacing	36	-0.05
(A) Compliance	Conforming	27	0.01	Cooperative	23	0.51	Individualistic	14	0.00	Uncooperative	49	-0.49
(A) Modesty	Self-deprecating	74	-0.16	Humble	17	0.41	Confident	34	0.08	Conceited	60	-0.45
(A) Tender-mindedness	Weak	56	-0.17	Sentimental	1	0.18	Tough	2	-0.10	Ruthless	40	-0.42
(C) Competence	Inept	60	-0.22	Novice	14	0.03	Competent	21	0.17	Know-it-all	42	-0.35
(C) Order	Rigid	48	-0.28	Orderly	27	0.26	Flexible	23	0.23	Chaotic	45	-0.30
(C) Dutifulness	Irresponsible	54	-0.37	Casual	8	0.17	Dutiful	15	0.30	Uptight	42	-0.21
(C) Achievement-striving	Unambitious	40	-0.25	Easy-going	19	0.35	Driven	21	0.23	Workaholic	0	0.06
(C) Self-discipline	Lazy	55	-0.31	Laid-back	9	0.18	Disciplined	23	0.29	Compulsive	35	-0.18
(C) Deliberation	Hasty	48	-0.33	Spontaneous	0	-0.13	Thoughtful	16	0.33	Over-controlled	49	-0.37
(N) Anxiety	Indifferent	50	-0.21	Relaxed	16	0.17	Alert	30	0.20	Tense	36	-0.15
(N) Angry-hostility	Timid	30	0.04	Agreeable	20	0.43	Strong-willed	17	0.04	Aggressive	8	-0.29
(N) Depression	Depressed	87	-0.34	Sensitive	0	0.19	Happy	37	0.40	Delirious	66	-0.25
(N) Self-conscious	Self-conscious	12	0.04	Self-aware	18	0.21	Natural	28	0.29	Superficial	95	-0.31
(N) Impulsiveness	Inhibited	49	-0.06	Self-controlled	17	0.27	Uninhibited	1	-0.17	Reckless	39	-0.41
(N) Vulnerability	Vulnerable	26	-0.02	Tender	4	0.25	Secure	62	0.22	Hard	34	-0.35
(O) Fantasy	Spaced-out	64	-0.28	Imaginative	16	0.04	Rational	22	0.29	Literal-minded	30	-0.10
(O) Aesthetics	Effete	67	-0.15	Sophisticated	22	0.11	Practical	27	0.23	Uncultured	46	-0.29
(O) Actions	Restricted	49	-0.18	Prudent	2	0.15	Adventurous	17	0.02	Wild	7	-0.25
(O) Feelings	Repressed	63	-0.19	Controlled	2	0.17	Expressive	13	0.08	Over-emotional	58	-0.23
(O) Ideas	Absent-minded	54	-0.20	Intellectual	14	0.15	Down-to-earth	38	0.34	Closed-minded	52	-0.18
(O) Liberalism	Radical	44	-0.33	Liberal	1	-0.09	Conservative	0	0.15	Authoritarian	0	-0.20
Mean		54	-0.21		13	0.16		22	0.17		43	-0.25

Note. E, extraversion; A, agreeableness; C, conscientiousness; N, neuroticism; O, openness to experience.

% Eval, percent of item communality attributable to evaluation.

SL, Schmid-Leiman *g* Loadings.



among raters, so we averaged the ratings into a mean evaluation score per item. These 120 ratings ranged from 1.25 for *depressed* to 4.78 for *happy*, and the distribution was bimodal, as is typical for evaluation ratings (Cruse, 1965). We then submitted the self-report data to a PC analysis, extracted the first PC, and correlated the loadings with the evaluation ratings. These correlated  $r=.93$ , replicating the finding that evaluation is the most important way in which participants respond to items (Edwards, 1969). Figure 1 displays this correlation, with overlaid distribution histograms.

### Exploratory factor analysis

We then conducted a conventional EFA of the self-report data. We analysed the data with MPLUS 5.2 (Muthén and Muthén, 3463 Stoner Avenue Los Angeles, CA 90066, 1998–2007), treating the responses as ordered categorical variables because of non-normal item distributions. That is, polychoric (rather than Pearson) correlations were modelled using unweighted least squares and a full-weight matrix. The first 10 Eigenvalues based on the polychoric correlation matrix were 16.21, 10.32, 8.74, 6.65, 4.23, 3.47, 2.37, 2.12, 1.92 and 1.85, indicating a substantial flattening of the scree plot after six factors. Subsequently, we extracted six factors and rotated the solution to simple structure using promax. The six-factor model fits the data reasonably well (root mean square error of approximation, RMSEA=0.06; root mean square residual=0.05). The items with the largest positive and negative loadings for each of the resulting factors are shown in Table 2, and the factor intercorrelations are displayed in Table 3. The first factor, which we labelled self-controlled versus reckless, showed loadings for *rational*, *co-operative* and *cautious* at one end, and for *reckless*, *wild* and *chaotic* at the other end. It appears to be a variant of conscientiousness. The second factor, which we called conservative versus liberal, showed positive loadings for *conservative*, *closed-minded* and *conforming*, and negative

loadings for *liberal*, *intellectual* and *sceptical*. It seems to represent a version of openness to experience and intellect. The third factor, which we called assertive versus weak, showed positive loadings for *strong-willed*, *assertive* and *confident*, and negative loadings for *weak*, *submissive* and *absent-minded*. It appears to represent a blend of extraversion and agreeableness. The fourth factor, which we labelled easy-going versus uptight, showed positive loadings for *laid-back*, *casual* and *flexible*, and negative loadings for *tense*, *over-controlled* and *workaholic*. It appears to represent a blend of neuroticism and conscientiousness. The fifth factor, which we labelled warm versus cold, showed positive loadings for *sensitive*, *tender* and *warm*, and negative loadings for *cold*, *ruthless* and *tough*. It appears to represent a blend of neuroticism and agreeableness. The sixth factor, which we labelled sociable-withdrawn, showed positive loadings for *happy*, *cheerful* and *expressive*, and negative loadings for *restricted*, *rigid* and *timid*. It appears to be a version of extraversion. As is usually the case with conventionally derived personality traits, these factors were quite evaluative such that one end contained primarily positively valenced markers, and the other end negatively valenced markers. To quantify this overlap, we correlated the loadings of each factor with the mean evaluation ratings and self-controlled correlated at  $r=.46$ , conservative at  $r=.05$ , assertive at  $r=.65$ , easy-going at  $r=.56$ , warm at  $r=.19$  and sociable at  $r=.48$ .

### Higher-order factor model

To examine how the evaluative dimension pertained to the higher-order factor of personality that has been proposed by Rushton et al. (2008), we extracted one higher-order factor by conducting a maximum-likelihood EFA of the factor intercorrelation matrix using the 'psych' package (Revelle, 2009) in R (R Development Core Team, 2010). The loadings on the higher-order factor consisted of self-control (loading=0.58), conservative (loading=0.33), assertive (loading=0.26), easy-going (loading=0.50), warm (loading=0.36) and sociable (loading=0.13). Thus, this higher-order factor appeared to dovetail with previous investigations in that it contrasted primarily socially desirable against primarily socially undesirable behaviours (Musek, 2007; Rushton et al., 2008).

### Hierarchical factor model

To investigate how items, rather than factors, loaded on this higher-order factor, we transformed the higher-order solution into a hierarchical solution by post-multiplying the original loading matrix with the higher-order loadings (Schmid & Leiman, 1957),

$$\Lambda_1 \Lambda_2$$

where  $\Lambda_1$  is the original promax factor loading matrix and  $\Lambda_2$  is the higher-order factor loading matrix.<sup>2</sup>

The loadings of the hierarchically transformed higher-order factor (displayed in Table 1) correlated  $r=.86$  with the mean

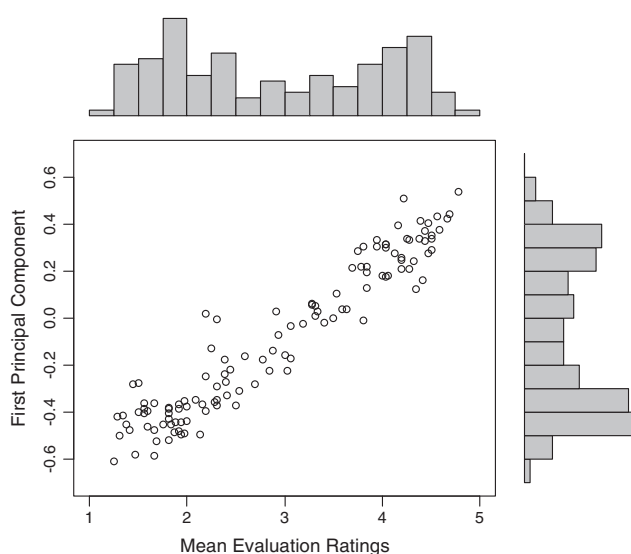


Figure 1. Correlation between mean evaluation ratings and the first principal component.

<sup>2</sup>As noted by an anonymous reviewer, this equation only derives the  $g$  loadings, not the subfactor loadings.

Table 2. Promax exploratory factor analysis loadings

Self-controlled–reckless		Conservative–liberal		Assertive–weak	
Item	Loading	Item	Loading	Item	Loading
Rational	0.31	Conservative	0.82	Strong-willed	0.66
Cooperative	0.30	Active	0.40	Assertive	0.63
Reserved	0.25	Disciplined	0.39	Driven	0.58
Self-controlled	0.25	Workaholic	0.29	Self-aware	0.56
Cautious	0.25	Closed-minded	0.28	Confident	0.55
Prudent	0.24	Happy	0.26	Rational	0.55
Orderly	0.24	Rigid	0.23	Secure	0.52
Practical	0.22	Conforming	0.23	Alert	0.52
Agreeable	0.20	Controlled	0.22	Competent	0.52
Discreet	0.19	Over-emotional	0.22	Authoritarian	0.51
Hasty	−0.54	Sceptical	−0.23	Spaced-out	−0.36
Aggressive	−0.55	Expressive	−0.23	Lazy	−0.38
Ruthless	−0.57	Sentimental	−0.24	Irresponsible	−0.38
Radical	−0.57	Intellectual	−0.25	Sluggish	−0.39
Manic	−0.62	Imaginative	−0.25	Gullible	−0.39
Spontaneous	−0.65	Sensitive	−0.26	Absent-minded	−0.40
Rash	−0.65	Depressed	−0.27	Inept	−0.40
Chaotic	−0.67	Sluggish	−0.28	Submissive	−0.44
Wild	−0.68	Lazy	−0.35	Weak	−0.51
Reckless	−0.76	Liberal	−0.83	Unambitious	−0.52

Easy-going–uptight		Warm–cold		Sociable–withdrawn	
Item	Loading	Item	Loading	Item	Loading
Easy-going	0.85	Sensitive	0.83	Happy	0.52
Laid-back	0.82	Sentimental	0.81	Sociable	0.51
Relaxed	0.72	Tender	0.72	Cheerful	0.47
Casual	0.60	Warm	0.66	Bold	0.36
Flexible	0.47	Overbearing	0.61	Confident	0.35
Humble	0.44	Gushy	0.58	Expressive	0.29
Modest	0.43	Vulnerable	0.56	Spontaneous	0.27
Down-to-earth	0.43	Expressive	0.52	Uninhibited	0.27
Agreeable	0.37	Gullible	0.45	Warm	0.26
Natural	0.37	Clingy	0.45	Indiscreet	0.26
Selfish	−0.26	Uncooperative	−0.19	Cautious	−0.50
Clingy	−0.26	Unambitious	−0.20	Discreet	−0.51
Workaholic	−0.29	Conservative	−0.21	Timid	−0.56
Conceited	−0.30	Laid-back	−0.21	Rigid	−0.56
Overbearing	−0.32	Relaxed	−0.22	Repressed	−0.60
Over-controlled	−0.32	Tough	−0.25	Withdrawn	−0.60
Uncooperative	−0.33	Ruthless	−0.26	Secretive	−0.61
Know-it-all	−0.36	Indifferent	−0.34	Restricted	−0.66
Tense	−0.40	Hard	−0.34	Private	−0.67
Uptight	−0.54	Cold	−0.41	Reserved	−0.76

Table 3. Promax factor intercorrelations

	Controlled vs reckless	Conservative vs liberal	Assertive vs Weak	Easy-going vs uptight	Warm vs cold	Sociable vs restricted
Controlled	1.00					
Conservative	0.17	1.00				
Assertive	0.19	0.11	1.00			
Easy-going	0.32	0.07	0.10	1.00		
Warm	0.17	0.33	0.03	0.18	1.00	
Sociable	0.04	−0.05	0.09	0.25	−0.13	1.00

evaluation ratings, suggesting that it represents a tendency to endorse positive versus negative self-descriptors. However, it may be that positively valenced items load in the same direction

because people who are extraverted are also emotionally stable, cooperative and so on, but given that the items were designed to allow for separation of valence from content, this hypothesis

could be examined in more detail. Interestingly, this dimension clustered items of opposite content as long as they had similar valence (e.g. *sluggish* and *manic*, *suspicious* and *gullible*, *grim* and *frivolous*, *secretive* and *indiscreet*, *rigid* and *chaotic*, *restricted* and *wild*, etc., all loaded in the same direction). As it seems unlikely that this dimension actually describes some consistent pattern of behavioural tendencies, it appears more plausible that it represents some kind of response bias.

### Isolating evaluation with exploratory structural equation modelling

We proceeded to isolate the evaluative factor onto a separate dimension orthogonal to the remaining factors using ESEM

(Asparouhov & Muthén, 2009). To identify the evaluation dimension, we calculated factor loadings that would produce the mean evaluation ratings as regression-based factor scoring weights using the formula

$$Rb = s$$

where  $R$  represents the item correlation matrix,  $b$  the centred mean evaluation ratings, and  $s$  the factor structure coefficients that are required. We fixed one factor to  $s$  and simultaneously extracted five additional exploratory factors orthogonal to this evaluative dimension. Thus, we essentially rotated the original factor matrix to one evaluation factor and five non-evaluative factors. As expected, this model fits

Table 4. Standardised exploratory structural equation modelling loadings based on target rotation

Evaluation		Extraversion		Agreeableness	
Item	Loading	Item	Loading	Item	Loading
Happy	0.59	Gushy	0.52	Gullible	0.35
Secure	0.54	Wild	0.51	Submissive	0.31
Cheerful	0.51	Sociable	0.49	Agreeable	0.30
Warm	0.50	Spontaneous	0.49	Trusting	0.28
Cooperative	0.50	Bold	0.47	Conservative	0.27
Sociable	0.47	Overbearing	0.46	Weak	0.23
Down-to-earth	0.47	Cheerful	0.46	Gushy	0.22
Confident	0.46	Expressive	0.43	Cheerful	0.22
Rational	0.44	Chaotic	0.41	Timid	0.21
Easy-going	0.44	Happy	0.40	Inept	0.21
Rigid	−0.50	Restricted	−0.28	Authoritarian	−0.38
Cold	−0.51	Timid	−0.29	Suspicious	−0.38
Superficial	−0.51	Rational	−0.30	Hard	−0.39
Uncooperative	−0.52	Secretive	−0.30	Tough	−0.39
Withdrawn	−0.52	Withdrawn	−0.32	Strong-willed	−0.40
Weak	−0.52	Cautious	−0.32	Self-reliant	−0.43
Selfish	−0.56	Placid	−0.34	Assertive	−0.43
Sluggish	−0.59	Discreet	−0.40	Sceptical	−0.47
Grim	−0.61	Private	−0.48	Frank	−0.49
Depressed	−0.63	Reserved	−0.61	Individualistic	−0.50
Conscientiousness		Neuroticism		Openness	
Item	Loading	Item	Loading	Item	Loading
Uptight	0.52	Sensitive	0.75	Liberal	0.34
Workaholic	0.46	Sentimental	0.69	Lazy	0.19
Driven	0.44	Tender	0.64	Expressive	0.19
Tense	0.42	Vulnerable	0.56	Know-it-all	0.17
Orderly	0.42	Warm	0.50	Intellectual	0.17
Serious	0.41	Overbearing	0.48	Selfish	0.14
Disciplined	0.40	Thoughtful	0.45	Competent	0.14
Dutiful	0.36	Timid	0.40	Conceited	0.11
Self-controlled	0.33	Clingy	0.39	Frank	0.10
Alert	0.29	Expressive	0.39	Uninhibited	0.10
Flexible	−0.34	Cold	−0.20	Reserved	−0.29
Sluggish	−0.38	Uncultured	−0.20	Repressed	−0.30
Irresponsible	−0.38	Authoritarian	−0.21	Over-emotional	−0.31
Absent-minded	−0.42	Bold	−0.21	Restricted	−0.35
Spaced-out	−0.45	Indiscreet	−0.22	Rigid	−0.36
Lazy	−0.46	Hard	−0.26	Modest	−0.38
Casual	−0.46	Tough	−0.26	Humble	−0.39
Relaxed	−0.59	Confident	−0.27	Conservative	−0.41
Easy-going	−0.63	Aggressive	−0.27	Disciplined	−0.41
Laid-back	−0.71	Ruthless	−0.27	Active	−0.42

the data about as well (RMSEA=0.05) as the aforementioned EFA.

The first factor (Table 4), which by design was fixed to evaluation, included positive loadings on *happy*, *cooperative* and *rational*, and negative loadings on *depressed*, *selfish* and *weak*. This dimension clustered items of opposite descriptive content but equal valence. For example, the evaluative factor comprised loadings of the same sign for *sluggish* (loading=−0.59) and *manic* (loading=−0.50), *grim* (loading=−0.61) and *frivolous* (loading=−0.37), *suspicious* (loading=−0.39) and *gullible* (loading=−0.24), and *self-deprecating* (loading=−0.43) and *conceited* (loading=−0.47). The evaluation factor also included positive loadings for pairs with opposite descriptive content, including *orderly* (loading=0.31) and *flexible* (loading=0.29), *modest* (loading=0.26) and *assertive* (loading=0.23), *easy-going* (loading=0.44) and *driven* (loading=0.40), and *intellectual* (loading=0.34) and *down-to-earth* (loading=0.47). Although one can imagine circumstances in which any one of these pairs might describe the actual behaviour of a respondent, it is unlikely that a behavioural style could be characterised by so many paradoxical pairs of descriptors, so this dimension may be better interpreted as a response bias rather than as a description of a consistent behavioural propensity.

The remaining exploratory factors were forced to be independent of the evaluation factor. Because the items were derived from the NEO PI-R (Costa & McCrae, 1992) facet structure, there were some preconceived notions about the multivariate structure. Therefore, we rotated the non-evaluative exploratory factors with an oblique target rotation described by Browne (2001). Conceptually, the target rotation may be viewed as sitting somewhere between an EFA and a CFA. It is useful when researchers have a theory-based expectation about the resulting latent structure. In contrast to CFA, loadings are not forced to zero but rather start at zero and may diverge away from zero if that better approximates the data. We created a target matrix by specifying that quadruples of items may only load on their appropriate factor for the Five-Factor Model (FFM), with other loadings starting at zero. For example, all but the 24 markers for the six extraversion facets were targeted to start at zero in order to rotate one factor as closely as possible to traditional extraversion as indicated by the NEO PI-R facet structure. The factor correlation matrix is displayed in Table 5. We also partitioned the total communality of each item into evaluative and descriptive variance, as displayed in Table 1. It is noteworthy that many of the negatively evaluated items contained almost no descriptive variation at all, which may indicate that the

prevalent problem of comorbidity among various forms of psychopathology (Grant, Stinson, Dawson, Chou, & Ruan, 2005; Krueger, 1999; Roysamb *et al.*, 2011) might be partly due to the social undesirability of descriptors.

Results suggested that the first non-evaluative dimension (Table 4), which was rotated to extraversion, had positive loadings not only on *spontaneous*, *sociable* and *expressive*, but also on items of negative valence such as *wild*, *gushy* and *overbearing*. It had negative loadings not only on *timid*, *withdrawn* and *restricted*, but also on items of more positive valence such as *cautious*, *private* and *discreet*. This structure was repeated for the rest of the non-evaluative factors: in place of the evaluative structure typical of the many personality models, with positively valenced items loading at one end and negatively valenced at the other, the factors included both positive and negative aspects of the relevant characteristics at both ends of the dimensions. The second non-evaluative factor, which was rotated to agreeableness, included positive loadings on *agreeable*, *trusting*, *submissive* and *weak*, and negative loadings on *authoritarian*, *hard*, *strong-willed* and *frank*. The third non-evaluative factor, which was rotated to conscientiousness, included positive loadings on *driven*, *disciplined*, *workaholic* and *uptight*, and negative loadings on *irresponsible*, *lazy*, *laid-back* and *flexible*. The fourth non-evaluative factor, which was rotated to neuroticism, loaded positively on *vulnerable*, *timid*, *sensitive* and *tender*, and negatively on *confident*, *tough*, *ruthless* and *cold*. The fifth non-evaluative dimension, which was rotated to openness, had positive loadings on *liberal*, *intellectual*, *know-it-all* and *conceited*, and negative loadings on *conservative*, *disciplined*, *rigid* and *restricted*. Thus, after isolating participants' tendency to endorse positive or negative markers—a tendency that in and of itself seems unlikely to describe a consistent behavioural pattern—the items clustered by their descriptive content, regardless of their valence. It is interesting to note that these factors overlap quite well with a factor solution based on adjectives with neutral social desirability values (Saucier, 1994), providing some evidence of convergent validity.

It should also be noted that the magnitude of the non-evaluative factor intercorrelations became quite small once evaluation was controlled for (*cf.* Bäckström, 2007; Bäckström *et al.*, 2009). The histogram of the factor intercorrelations without isolating evaluation (from Table 3) are displayed on the left side of Figure 2, and the factor intercorrelations among the non-evaluative factors (from Table 5) are displayed on the right side. Although extracting a GFP from the non-evaluative correlation matrix is feasible, it would

Table 5. Target oblique factor intercorrelations

	Positive vs negative evaluation	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness to experience
Positive evaluation	1.00					
Extraversion	0.00	1.00				
Agreeableness	0.00	−0.14	1.00			
Conscientiousness	0.00	−0.11	−0.17	1.00		
Neuroticism	0.00	−0.10	0.11	0.02	1.00	
Openness to experience	0.00	0.03	0.08	−0.08	−0.21	1.00



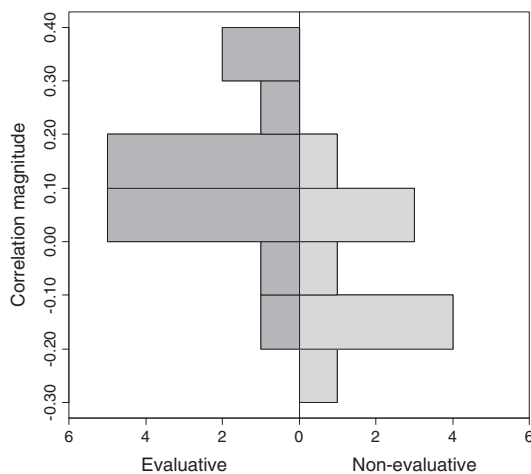


Figure 2. Histogram of factor correlations with evaluation included (left side) and excluded (right side).

not be particularly meaningful as there is so little covariation to account for. In other words, the GFP appears to capitalise on covariation stemming from shared valence among items.

## DISCUSSION

In this study, we developed a set of personality trait markers inspired by Peabody (1967). These trait markers allowed for separating the variance attributable to valence and description into two separate sources. After analysing self-report responses to these items, it was demonstrated that the GFP correlated almost perfectly with the evaluation ratings of the items. Furthermore, it was observed that this dimension clustered items of opposite content as long as both constructs had similar valence (e.g. *manic* and *sluggish* loaded in the same direction), rendering it an unlikely description of a consistent behavioural pattern. Isolating this evaluative dimension rendered a solution in which items are clustered by their descriptive properties (e.g. *active* and *manic* loaded in the same direction). Finally, the covariation among the non-evaluative factors was quite small, implying that the GFP capitalises on evaluative rather than descriptive covariance.

Rushton et al. (2008) argued that individuals who exhibited a greater degree of socially desirable traits, such as extraversion and emotional stability, had a selective advantage, and individuals who were introverted, neurotic, and so forth, were gradually weeded out. They proposed that 'those at the high end of the GFP likely enjoyed greater reproductive success' and that 'unidirectional selection toward more efficient and cooperative personalities was a driving force leading from the archaic hominid to the modern personality' (pp. 1181). In other words, they hypothesised that the relationship between the GFP and reproductive success is mediated by behaviour characterised by extraversion, emotional stability, conscientiousness, agreeableness and openness to experience, but as we have demonstrated, people high (or low) on this dimension endorsed constructs of opposite meaning as long

as they were positively (or negatively) evaluated in terms of social norms. Could it be that a biologically based trait shaped individuals to act both placid and active, both laid-back and disciplined, both agreeable and strong-willed, and so on? Although the exact nature of this dimension remains to be explored, the fact that it ignores item content weakens the argument that the GFP evolved to generate a consistent behavioural style that in turn afforded greater evolutionary success.

An alternate interpretation is that the GFP taps general self-esteem, such that people high on this trait feel so good that they endorse any positively valenced trait, no matter what the item is actually designed to measure. It might also measure general cultural knowledge of which words have positive versus negative valence.<sup>3</sup> That is, the primary difference between someone who endorses *active* versus *manic* is not necessarily a behavioural one, but rather an awareness of cultural norms. Another interpretation is that the GFP represents some kind of response bias (Paulhus, 1984). That is, this dimension may not have the typical correspondence to a particular cluster of feelings, thoughts and behaviours that one usually attributes to traits. For example, if individuals endorse items pertaining to sociability and cheerfulness, we tend to presume that they are more likely than the average person to strike up congenial conversations and frequent social gatherings, but such a deduction simply does not follow from the evaluation dimension.

Rushton et al. (2008), however, have argued against interpreting the GFP as a response bias. By regressing out scores on social desirability inventories and still finding support for the GFP, they propose that the GFP represents more than a tendency to endorse items based on their valence. Erdle and Rushton (2010) analysed self-report responses to personality scales and found that the relation among these traits and the first PC, which they argued served as a proxy for the GFP, remained virtually unchanged after the Eysenck Lie scale was controlled for. We remain sceptical, though, that typical social desirability or lie scales capture the valence of items. In one study, the correlations between a lie scale and the GFP were 0.22 and 0.25 in two separate samples, respectively (Schermer & Vernon, 2010). Although these relations are not trivial, they are far from the correlation of 0.86 between the GFP and the item evaluation ratings that we obtained in this sample. Therefore, we do not think that regressing out contemporary lie or social desirability scales controls for evaluation very well. More powerful approaches to remove evaluation include using ipsatised items (Hirsh & Peterson, 2008), neutral markers (Bäckström et al., 2009; Saucier, 1994), statistical techniques such as target rotation (Peabody & Goldberg, 1989) or ESEM (Asparouhov & Muthén, 2009), or cross-rater designs (Anusic et al., 2009; DeYoung, 2006; McCrae et al., 2008). Moreover, when such approaches are applied, the factor intercorrelations decrease substantially, diminishing the importance of the GFP.

<sup>3</sup>We thank an anonymous reviewer for pointing this out.

## General factor of personality and evolutionary views of personality

The view of Rushton *et al.* (2008) that personality is an adaptation with only one correct answer stands in contrast to the structure of the non-evaluative factors. These factors, with positively and negatively valenced markers at each end of the factors, suggest that different kinds of personality constellations proffer adaptive value but possibly in different contexts. Recently, Penke *et al.* (2007) argued that higher intelligence is monotonically related to greater adaptive value, whereas higher scores on personality traits in socially desirable directions do not always offer greater adaptive value. They suggested that differences in traits across individuals evolved because different environments afforded different survival strategies. Likewise, Nettle (2006) argued that both high and low scores on Big Five traits may have evolutionary advantages.

The results based on the non-evaluative factors align closely with such theories. For example, in our sample, extraversion was marked not only by positively valenced markers such as *sociable* but also by negatively valenced ones such as *overbearing*. Introverts, on the other hand, were *withdrawn* but also *discreet*. Emotionally stable individuals were not only *tough* but also *cold*, whereas neurotic individuals were not only *vulnerable* (as they are typically described by most personality inventories) but also *sensitive*. Additionally, it seems relatively easy to imagine environments or situations where traits like sensitivity or discreteness would be adaptive and traits such as coldness or arrogance would be maladaptive. In other words, the non-evaluative factors appear to represent alternate adaptive solutions to ongoing problems in living (cf. McCrae, Löckenhoff, & Costa, 2005). A person can be more or less sensitive to unpleasant external stimuli, and at moderate levels, either orientation may be adaptive and positively evaluated by others. At extreme levels of either sensitivity or insensitivity, both become negatively evaluated, and one presumes maladaptive in their consequences.

## Implications for personality measurement

Although we primarily address the GFP in this paper, we believe that our results also have implications for personality measurement as a whole. We conjecture that there are at least two independent sources of variance in self-reported personality: a face-valid portion representing respondents' effort to describe themselves in terms of the semantic content of the items, and an evaluative portion representing their tendency to describe themselves either positively or negatively with regards to social norms, regardless of the semantic content of the items. Common personality models rotate the two sources of response bias into a set of amalgam dimensions that are both descriptive and evaluative, usually in unknown proportions. What happens to evaluation in most personality models, such as the Big Five? The evaluation dimension is just as prominent in a five-factor inventory as it is in any other set of personality items, but in the FFM, it is rotated so that its variation is incorporated in the resulting five dimensions. Therefore, the familiar dimensions of the FFM

are themselves highly evaluative. In most Big Five inventories, there are no positively evaluated ways in which one could be disagreeable or neurotic. As a consequence, it may be difficult to know whether individuals scoring at the high (or low) end of such dimensions primarily endorsed positively evaluated items or whether they responded to the actual content of the items irrespective of their perceived valence. Therefore, we think that there is some theoretical merit in rotating solutions to an evaluation dimension and several non-evaluative dimensions in order to bring clarity to respondents' scale scores.

As a final note, it is important to note that we were able to conduct the analyses that we reported here because our factor analyses were conducted at the item level. Most investigations of higher-order structures, however, are usually conducted at the scale level. Although evaluation and description content could also be separated at the scale level, this is a more difficult task because the information regarding how individuals respond to items is lost during scale aggregation.

## CONCLUSION

The GFP has surfaced in different shapes since the 1950s. The most recent interpretation attributes evolutionary importance to it, proposing that people who were extraverted, agreeable, hard-working, emotionally stable and open to experience experienced more success at survival and reproduction, but caveats regarding this interpretation exist. First, when items of a neutral kind are analysed, the GFP all but disappears. Second, when observer biases are taken into account, the GFP disappears. Third, as we have demonstrated, the GFP clusters items of opposite meaning as long as they have similar valence. Although it is important to know whether individuals primarily endorse items of positive versus negative valence, it does not seem reasonable to consider this dimension as a trait that corresponds to a coherent behavioural style. A more plausible interpretation is that it represents some kind of response bias. In addition, isolating this bias renders a factor solution that appears to better approximate evolutionary perspectives of personality as a specialisation for a particular environmental niche. Whereas Rushton *et al.* (2008) concluded by suggesting that all efficient personalities are alike, we believe successful personalities come in many different shapes and forms.

## REFERENCES

- Amigo, S., Caselles, A., & Mico, J. C. (2010). General factor of personality questionnaire (GFPQ): Only one factor to understand personality? *The Spanish Journal of Psychology*, 13, 5–17.
- Anusic, I., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2009). The nature and structure of correlations among Big Five ratings: The Halo-Alpha-Beta model. *Journal of Personality and Social Psychology*, 97, 1142–1156.
- Ashton, M. C., Lee, K., Goldberg, L. R., & de Vries, R. E. (2009). Higher-order factors of personality: Do they exist? *Personality and Social Psychology Review*, 13, 79–91.
- Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, 16, 397–438.

- Bäckström, M. (2007). Higher-order factors in a five-factor personality inventory and its relation to social desirability. *European Journal of Psychological Assessment*, 23, 63–70.
- Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, 43, 335–344.
- Block, J. (1965). *The challenge of response sets*. East Norwalk, CT: Appleton-Century-Crofts.
- Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36, 111–150.
- Costa, P. T. Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory and NEO Five-Factor Inventory. Professional Manual. Psychological Assessment Resources, Inc. Odessa, FL.
- Crowne, D. P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence*. New York: Wiley.
- Cruse, D. B. (1965). Social desirability scale values of personal concepts. *Journal of Applied Psychology*, 49, 342–344.
- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, 91, 1138–1151.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden Press.
- Edwards, A. L. (1969). Trait and evaluative consistency in self-description. *Educational and Psychological Measurement*, 29, 737–752.
- Erdle, S., & Rushton, J. P. (2010). The general factor of personality, BIS–BAS, expectancies of reward and punishment, self-esteem, and positive and negative affect. *Personality and Individual Differences*, 48, 762–766.
- Figueredo, A. J., & Rushton, J. P. (2009). Evidence for shared genetic dominance between the general factor of personality, mental and physical health, and life history traits. *Twin Research and Human Genetics*, 12, 555–563.
- Grant, B. F., Stinson, F. S., Dawson, D. A., Chou, S. P., & Ruan, W. J. (2005). Co-occurrence of DSM-IV personality disorders in the United States: Results from the National Epidemiological Survey on Alcohol and Related Conditions. *Comprehensive Psychiatry*, 46, 1–5.
- Heilbrun, A. B. (1964). Social-learning theory, social desirability, and the MMPI. *Psychological Bulletin*, 61, 377–387.
- Hirsh, J. B., & Peterson, J. B. (2008). Predicting creativity and academic success with a 'fake-proof' measure of the Big Five. *Journal of Research in Personality*, 42, 1323–1333.
- Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, 56, 921–926.
- van der Linden, D., Scholte, R. H. J., Cillessen, A. H. N., te Nijenhuis, J., & Segers, E. (2010). Classroom ratings of likeability and popularity are related to the Big Five and the general factor of personality. *Journal of Research in Personality*, 44, 669–672.
- van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, 44, 315–327.
- Loehlin, J. C., & Martin, N. G. (2011). The general factor of personality: Questions and elaborations. *Journal of Research in Personality*, 45, 44–49.
- McCrae, R. R., Lökenhoff, C. E., & Costa, P. T. Jr. (2005). A step toward DSM-V: Cataloguing personality-related problems in living. *European Journal of Personality*, 19, 269–286.
- McCrae, R. R., Yamagata, S., Jang, K. L., Riemann, R., Ando, J., Ono, Y., et al. (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 95, 442–455.
- Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, 41, 1213–1233.
- Muthén, L. K., & Muthén, B. O. (1998–2007). *Mplus User's Guide* (5th Ed.). Los Angeles, CA: Muthén & Muthén
- Nettle, D. (2006). The evolution of personality variation in humans and other variables. *American Psychologist*, 61, 622–631.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598–609.
- Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology*, 7, 1–18.
- Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology*, 57, 552–567.
- Penke, L., Denissen, J. J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality*, 21, 549–587.
- R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Revelle, W. (2009). psych: Procedures for Personality and Psychological Research. R package version 1.0-70 [Computer software]. Retrieved February 2, 2010, from <http://www.r-project.org/>
- Roysamb, E., Kendler, K. S., Tambs, K., Orstavik, R. E., Neale, M. C., Aggen, S. H.,...Reichborn-Kjennerud, T. (2011). The joint structure of DSM-IV Axis I and Axis II disorders. *Journal of Abnormal Psychology*, 120, 198–209.
- Rushton, J. P., & Irwing, P. (2008). A general factor of personality (GFP) from two meta-analyses of the Big Five: Digman (1997) and Mount, Barrick, Scullen, and Rounds (2005). *Personality and Individual Differences*, 45, 679–683.
- Rushton, J. P., & Irwing, P. (2009a). A general factor of personality in the Comrey Personality Scales, the Minnesota Multiphasic Personality Inventory-2, and the Multicultural Personality Questionnaire. *Personality and Individual Differences*, 46, 437–442.
- Rushton, J. P., & Irwing, P. (2009b). A general factor of personality in the Millon Clinical Multiaxial Inventory-III, the Dimensional Assessment of Personality Pathology, and the Personality Assessment Inventory. *Journal of Research in Personality*, 43, 1091–1095.
- Rushton, J. P., & Irwing, P. (2009c). A general factor of personality in 16 sets of the Big Five, the Guilford–Zimmerman Temperament Survey, the California Psychological Inventory, and the Temperament and Character Inventory. *Personality and Individual Differences*, 47, 558–564.
- Rushton, J. P., Bons, T. A., Ando, J., Hur, Y., Irwing, P., Vernon, P. A.,...Barbaranelli, C. (2009). A general factor of personality from multitrait-multimethod data and cross-national twins. *Twin Research and Human Genetics*, 12, 356–365.
- Rushton, J. P., Bons, T. A., & Hur, Y. (2008). The genetics and evolution of the general factor of personality. *Journal of Research in Personality*, 42, 1173–1185.
- Saucier, G. (1994). Separating description and evaluation in the structure of personality attributes. *Journal of Personality and Social Psychology*, 66, 141–154.
- Schermer, J. A., & Vernon, P. A. (2010). The correlation between general intelligence (g), a general factor of personality (GFP), and social desirability. *Personality and Individual Differences*, 48, 187–189.
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53–61.
- Veselka, L., Schermer, J. A., Petrides, K. V., & Vernon, P. A. (2009). Evidence for a heritable general factor of personality in two studies. *Twin Research and Human Genetics*, 12, 254–260.
- de Vries, R. E. (2011a). No evidence for a general factor of personality in the HEXACO personality inventory. *Personality and Individual Differences*, 45, 229–232.
- de Vries, R. E. (2011b). No support for a general factor of personality in a reanalysis of Van der Linden et al. (2010). *Personality and Individual Differences*, 50, 512–516.
- Welsh, G. S. (1956). Factor dimensions A and R. In G. S. Welsh, & W. D. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine* (pp. 264–281). Minneapolis: University of Minnesota Press.