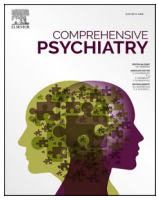
Comprehensive Psychiatry 134 (2024) 152514

Contents lists available at ScienceDirect

Comprehensive Psychiatry

journal homepage: www.elsevier.com/locate/comppsych



Validation of an abbreviated Big Five personality inventory at large population scale: Psychometric structure and associations with common psychiatric and neurological disorders

Weixi Kang [a,b,*], Jeggan Tiego [c], Peter J. Hellyer [b], William Trender [b], Jon E. Grant [d], Samuel R. Chamberlain [e,f], Adam Hampshire [b]

[a] School of Arts and Humanities, Tung Wah College, Hong Kong
[b] Department of Brain Sciences, Imperial College London, UK
[c] Neural Systems and Behavior Laboratory, The Turner Institute for Brain and Mental Health, School of Psychological Sciences and Monash Biomedical Imaging Facility, Monash University, Australia
[d] Department of Psychiatry, University of Chicago, Chicago, USA
[e] Department of Psychiatry, Faculty of Medicine, University of Southampton, Southampton, UK
[f] NHS Southern Gambling Service, and Southern Health NHS Foundation Trust, Southampton, UK.

ARTICLE INFO

Keywords:
Personality
Big Five
Validation
Reliability
Psychopathology
neurology

ABSTRACT

Background: The five-factor model of personality, as quantified using instruments such as the Big Five Inventory, consists of broad personality domains including Extraversion, Agreeableness, Conscientiousness, Neuroticism (emotional instability), and Openness. Such instruments typically include >40 items. However, instruments with many items can be unwieldly and a cause of measurement error in clinical and cohort studies where multiple scales are sequenced. Conversely, established 5- and 10-item versions of the Big Five Inventory have poor reliability. Here, we developed and validated an abbreviated 18-item Big Five Inventory that balances efficiency, reliability and sensitivity.

Method: We analysed three datasets (N = 59,797, N = 21,177, and N = 87,983) from individuals who participated in the online Great British Intelligence Test (GBIT) study, a collaborative citizen science project with BBC2 Horizon. We applied factor analyses (FA), predictive normative modelling, and one-sample t-tests to validate the 18-item version of the Big Five and to investigate its associations with psychiatric and neurological conditions.

Results: The 18-item version of the Big Five Inventory had higher validity and retest reliability compared to the other previously shortened versions in the literature, with comparable demographic associations to the full Big Five Inventory. It exhibited strong (i.e. large effect size) associations with psychiatric conditions, and moderate (small-medium) associations with neurological conditions. Neuroticism (emotional instability) was substantially higher in all psychiatric conditions, whereas Conscientiousness, Openness and Extraversion showed differential associations across conditions.

Conclusion: The newly validated 18-item version of the Big Five provides a convenient means of measuring personality traits that is suitable for deployment in a range of studies. It retains psychometric structure, retest reliability and clinical-group sensitivity, as compared to the full original scale.

1. Introduction

The five-factor model of personality is a set of broad personality domains consisting of Extraversion, Agreeableness, Conscientiousness, Neuroticism (emotional instability), and Openness (i.e. 'the Big Five') [1]. As opposed to being reserved and quiet, individuals who score high in Extraversion are assertive and sociable. People scoring high on Agreeableness are polite and cooperative, rather than rude and antagonistic. Individuals who score high in Conscientiousness are orderly and task-focused, rather than being disorganized and distractible. People

* Corresponding author at: School of Arts and Humanities, Tung Wah College Cheung Kung Hai Memorial Building, 90A Shantung Street, Mong Kok, Kowloon, Hong Kong.
E-mail address: wkang@twc.edu.hk (W. Kang).

https://doi.org/10.1016/j.comppsych.2024.152514

Available online 29 June 2024

Table 1
Selected personality items based on the heaviest loadings.

…someone who is talkative.
…someone who does a thorough job.
…someone who is original, comes up with new ideas.
…someone who is reserved.
…someone who is relaxed, handles stress well.
…someone who has a forgiving nature.
…someone who tends to be disorganized.
…someone who worries a lot.
…someone who tends to be quiet.
…someone who is emotionally stable, not easily upset.
...someone who is inventive.
...someone who perseveres until the task is finished.
…someone who values artistic, aesthetic experiences.
…someone who is considerate and kind to almost everyone.
…someone who likes to reflect, play with ideas.
…someone who has few artistic interests.
…someone who likes to cooperate with others.
…someone who is sophisticated in art, music, or literature.

Table 2
Demographic information of the participants of Dataset 3.

Variables	Mean	S.D.
Age	46.58	15.84
	Count	%
Handedness		
Right-handed	76,226	86.64
Left-handed	9480	10.77
Ambidextrous	2277	2.59
Gender		
Male	48,717	55.37
Female	38,876	44.19
Other	390	0.44
Language		
English	82,371	93.62
Other	5612	6.38
Ethnicity		
White European or North American	80,456	91.44
Other	7527	8.56
Country		
United Kingdom	80,539	91.54
Other	7444	8.46
Education		
University degree	51,088	58.07
Secondary school/High school diploma	31,406	35.70
PhD	3595	4.09
Primary/Elementary school	1772	2.01
No schooling	122	0.14
Income		
£10-20 K	9175	10.43
£20-30 K	11,806	13.42
£30-40 K	10,513	11.95
£40-50 K	7413	8.43
£50-60 K	4399	5.00
£60-70 K	2598	2.95
£70-80 K	1877	2.13
£80-90 K	1192	1.35
£90-100 K	1143	1.30
>100 K	3395	3.86
Unknown	31,358	35.64
Prefer not to say	2107	2.39

Table 3
The correlation coefficients (r) between scores from the full-scale Big Five and the shorter versions of the Big Five.

Trait	18-item	10-item	5-item
Openness	0.95	0.92	0.86
Extraversion	0.94	0.90	0.80
Neuroticism	0.91	0.84	0.71
Agreeableness	0.94	0.75	0.76
Conscientiousness	0.86	0.72	0.72

scoring high on Neuroticism (emotional instability) tend to experience negative emotions rather than being emotionally resilient. Finally, people who score high in Openness have broad interests, are sensitive to art and beauty, prefer novelty rather than having narrow interests, being indifferent to art or beauty and preferring routine. It is agreed in the literature that the Big Five capture some of the most important fundamental individual differences in personality traits. Moreover, many other personality models can be conceptualised in terms of the Big Five [2].

Various scales have been developed to measure the Big Five personality traits, with the Big Five Inventory (BFI) being amongst the most commonly applied. However, instruments for the Big Five personality assessments are typically very long, e.g., the BFI comprises >40 items [3]. Empirical evidence indicates that questionnaire scales with many items are sometimes a source of measurement error [4]. For instance, longer scales induce discouragement, fatigue, and inattention in participants, making them hard to deploy, particularly in clinical populations; and can lead to bias plus missing data due to non-completion. This issue is exacerbated in research studies that combine multiple established scales in sequence, which is common in clinical cohort, epidemiological and citizen science studies. Therefore, scales with fewer items are favoured if they achieve acceptable reliability and validity [4]. Consequently, a research objective has been to develop shortened scales (e.g., [5–8]). However, abbreviated scales are not always validated in large populations with cross-validation, which put them at risk of generalising poorly and providing unreliable estimates [9,10]. Although some studies have validated shortened Big Five [7,11,12], there is still a need for a short but rigorously validated personality inventory suitable for use in different contexts including via online platforms and applying to patients who may have too much cognitive burden if taking the original inventory.

Common psychiatric and neurological disorders are known to be associated with personality traits. Four alternative models have been proposed to account for the empirically demonstrated associations between personality and psychopathology; the predisposition/vulnerability model, complication/scar model, pathoplasty/exacerbation model and the spectrum model [13–16]. The predisposition/vulnerability model suggests that particular traits confer a greater risk for the development of a psychological disorder [13,15]. Conversely, the complication/scar model proposes that the development of psychological disorders alters an individual's premorbid personality [15,16]. According to the pathoplasty/exacerbation model, the etiology of psychological disorders is independent of traits, however, premorbid temperament traits are purported to influence the manifestation, severity, and course of mental illness [15,16]. The spectrum model conceptualises a dimensional relationship between temperament (i.e., a subset of personality)/personality and psychopathology, such that

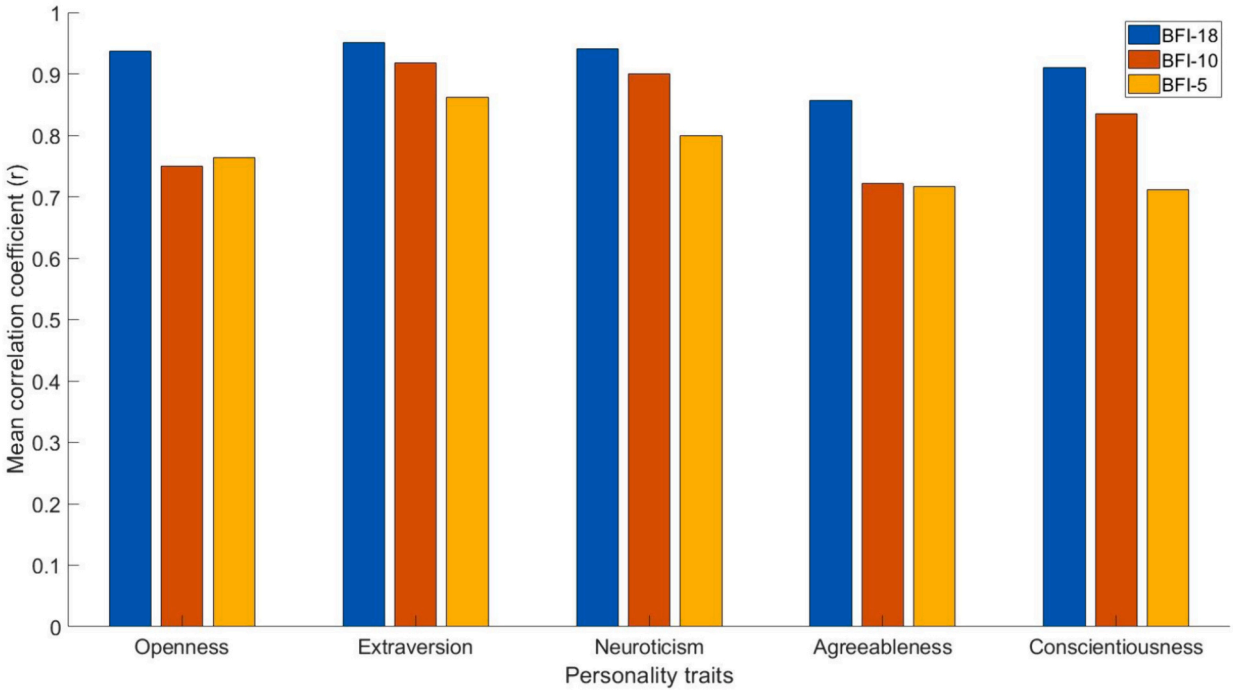


Fig. 1. Correlation between scores from the full-scale Big Five and the shorter versions of the Big Five. Factor scores for the 18, 10 and 5 item versions of the Big 5 were calculated and correlated against the full 44 item inventory. Y axes = mean correlation coefficient (r).

Table 4
Retest correlation coefficients (r) for 18-item Big Five.

Trait	r
Openness	0.84
Extraversion	0.83
Neuroticism	0.81
Agreeableness	0.81
Conscientiousness	0.73

psychological disorders represent extreme poles of the continuum of individual differences in temperament [13–15]. Indeed, empirical evidence found that Neuroticism is the trait that is most consistently associated with psychiatric conditions although other traits are also related to psychiatric and neurological conditions [17–27]. However, neurological and psychiatric groups have not been compared comprehensively side by side within the context of the same study at a large scale, where demographic factors can be accounted for.

The aim of this study was to determine the validity of a shortened BFI using three large-scale datasets. First, we identified a subset of items that represent different factors as defined by the full original BFI. Next, we evaluated the correspondence of the resultant factor scores across full and abbreviated scales and examined their retest reliabilities. Finally, we tested the associations between the factor scores from the abbreviated BFI and demographic variables, and psychiatric and neurological conditions.

2. Methods

2.1. Participants

All data were from participants in the online Great British Intelligence Test (GBIT) study, a citizen science project that was run in collaboration with BBC2 Horizon, which was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2008. All procedures were approved by the Imperial College Research Ethics Committee (17IC4009). The data were collected at three timepoints. All participants provided informed consent prior to completing the survey. Participants younger than 16 years old, older than 99 or who had missing data were removed from the datasets. Dataset 1 (N = 59,797) was collected from January through April 2020 and contains the full version of the Big Five (44 items). Dataset 2 (N = 21,177) was collected in December 2020 through recontact of participants from Dataset 1. Dataset 3 (N = 87,983) was collected was collected from a further independent cohort from May 2020 – December 2020 and included the 18-item version of the Big Five.

2.2. Materials

Participants were asked to complete a series of cognitive tasks and questionnaires online in the GBIT (Great Britain Intelligence Test) study. For analysis, we extracted age, handedness (right handed vs. left handed vs ambidextrous), gender (female vs. males vs. other), language (English vs. other), ethnicity (white European or North American vs. other), country of residence (United Kingdom vs. other), education (PhD vs. university degree vs. secondary school vs. high school diploma vs. primary/elementary school vs. no schooling), occupation (worker vs. retired vs. student vs. disabled/not applicable/sheltered employment vs. unemployed/looking for work vs. homemaker vs. unknown), self-report of previous formal diagnosis from a healthcare professional of psychiatric conditions (depression vs. anxiety vs. depression + anxiety vs. ADHD vs. OCD vs. bipolar vs. other) and neurological conditions (learning disability vs. multiple sclerosis vs. Parkinson's disease vs. traumatic brain injury vs. other). Depending on timepoint, we also collected either the full 44-item version of the BFI (supplementary material Table 1), or an abbreviated version consisting of 18 items (Table 1) that were selected based on factor analysis showing them to have strong loadings on specific personality traits (please refer to https://www.dropbox.com/scl/fi/f4bgezboq0f7wt9ph06m0/5KQselection.xlsx?rlkey=xk6fs0exzhq67r8gj1zx10jkv&dl=0 for more information).

2.3. Data analysis

All analyses were performed in MATLAB 2018a.

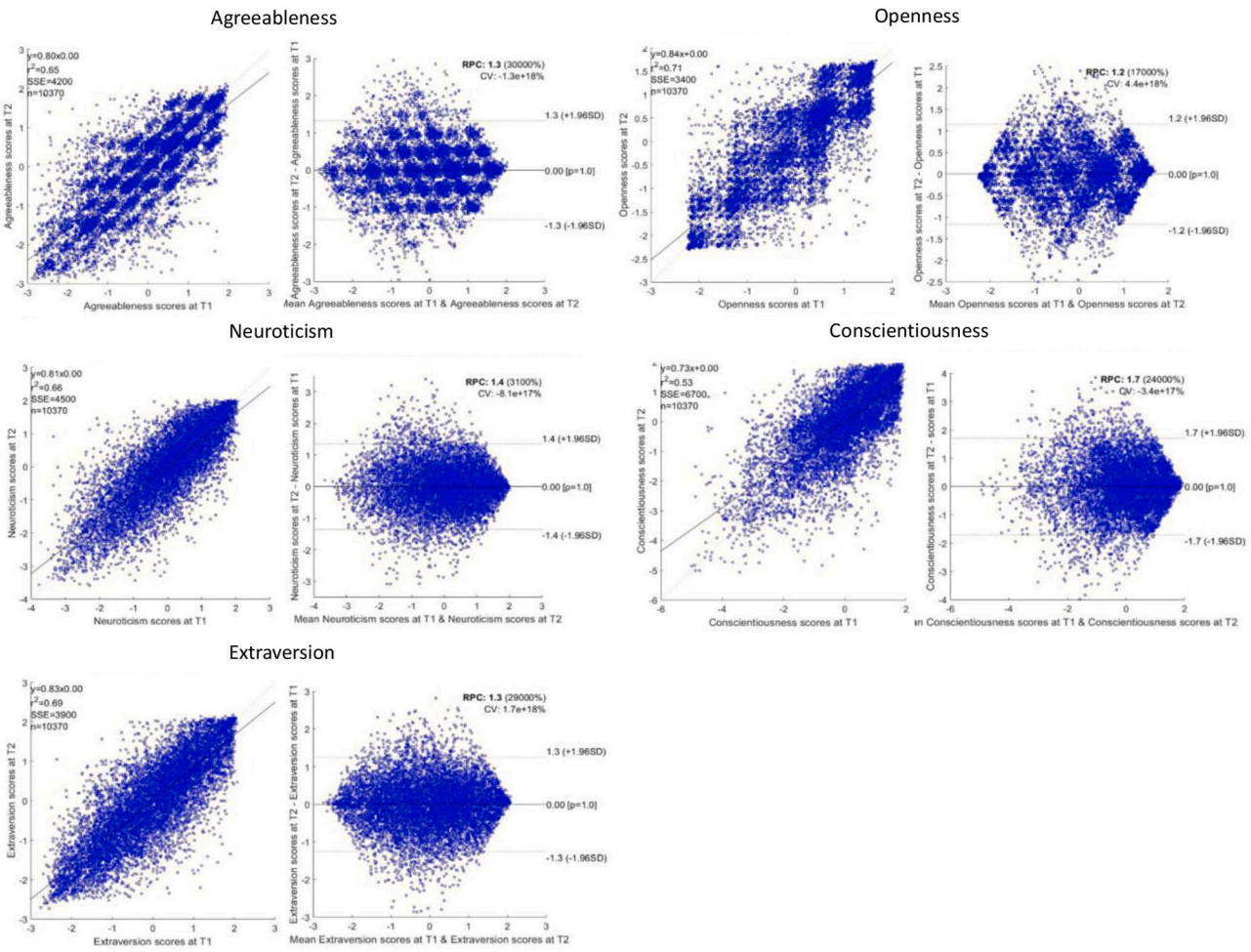


Fig. 2. Retest correlation and Bland–Altman plots for 18-item Big Five. Openness, Extraversion, Neuroticism (emotional instability), Agreeableness, and Conscientiousness scores were compared within subject across Time 1 and Time 2. Left panel: R^2 = Pearson r value squared, SSE = sum of squared error, and n = number of data points used. Right panel: RPC = reproducibility coefficient (1.96*SD), CV = coefficient of variation (SD of mean values in %).

2.4. Correspondence of full and abbreviated scales

We first analysed the full-scale Big Five Personality Inventory collected from GBIT Dataset 1 (N = 87,983). These data were randomly split into two halves. An exploratory factor analysis was run on one of these halves with 5 oblique components, producing an item-factor loading matrix (supplementary material Table 1). Factor scores were then estimated for the remaining half of the participants by regressing the factor loadings onto the participant scores. This was done for (i) the full set of items, (ii) the 18 items of the Abbreviated Big Five, (iii) the 10-item version of the Big Five proposed by Rammstedt et al. [7] (supplementary material Table 2), and (iv) the 5-item version selected based on the 5 most heavily loaded questions for each factor (i.e., personality trait; supplementary material Table 3). The strength of the correlation between the full item and reduced item scales were estimated to gauge conformity.

2.5. Retest reliability

Dataset 1 and Dataset 2 were analysed to confirm retest reliability across timepoints. Exploratory factor analyses were performed independently on Dataset 1 and Dataset 2 with the number of factors set to 5. Promax rotated item-factor loadings were examined to confirm that the resultant factors were interpretable as Openness, Extraversion, Neuroticism (emotional instability), Agreeableness, and Conscientiousness. The factor scores for each personality trait were then correlated within-participant across timepoints to quantify retest reliability.

2.6. Associations with psychiatric and neurological conditions

To determine how the Abbreviated Big Five is related to psychiatric and neurological conditions, we applied a predictive normative modelling approach to Dataset 3 as below.

(1) Normative models were trained to predict each factor from demographic variables including age, gender, handedness, language, ethnicity, country, education, occupation, and income in people without psychiatric or neurological conditions respectively by contracting.
(2) The normative models were used to transform the factor scores of participants who had psychiatric or neurological conditions into standard deviation from expected units (DfE), that is, the number of standard deviations by which each participant differed from what would be expected for people who have the same socio-demographic profiles.
(3) *t*-tests were performed to compare the differences in the DfE scores (comparable to effect sizes in Cohen's d) against zero.

This approach offers greater advantages compared to paired-sample t-tests because it can consider demographic variables that might impact cognitive abilities and handle imbalanced sample sizes. We controlled for these demographic factors because age, gender, education,

Table 5
Results from linear models with demographic predictors of personality traits including a. Openness, b. Extraversion, c. Neuroticism (emotional instability), d. Agreeableness, and e. Conscientiousness.

a. Openness

Variable	SumSq	DF	MeanSq	F	pValue
age	158.29	1	158.29	171.42	<0.001
handedness	106.57	2	53.29	57.71	<0.001
gender	635.66	2	317.83	344.21	<0.002
language	76.42	1	76.42	82.77	<0.001
ethnicity	2.065	1	2.065	2.24	0.135
country	43.01	1	43.01	46.58	<0.001
education	648.49	4	162.12	175.58	<0.001
income	24.81	11	2.26	2.44	0.005
occupation	276.76	5	55.35	59.95	<0.001
Error	63,938	69,244	0.92		

b. Extraversion

Variable	SumSq	DF	MeanSq	F	pValue
age	21.55	1	21.55	22.37	<0.001
handedness	2.28	2	1.14	1.18	0.307
gender	1142.70	2	571.37	593.14	<0.001
language	0.54	1	0.54	0.56	0.455
ethnicity	1.75	1	1.75	1.81	0.178
country	0.01	1	0.01	0.01	0.938
education	132.27	4	33.07	34.33	<0.001
income	540.51	11	49.14	51.01	<0.001
occupation	15.52	5	3.10	3.22	0.007
Error	66,704	69,244	0.96		

c. Neuroticism (emotional instability)

Variable	SumSq	DF	MeanSq	F	pValue
age	237.87	1	237.87	299.02	<0.001
handedness	7.08	2	3.54	4.45	0.012
gender	2254.20	2	1127.10	1416.90	<0.001
language	45.09	1	45.09	56.69	<0.001
ethnicity	5.23	1	5.23	6.57	0.010
country	0.024	1	0.024	0.030	0.862
education	31.12	4	7.78	9.78	<0.001
income	548.31	11	49.85	62.66	<0.001
occupation	87.976	5	17.595	22.119	3.1989e-22
Error	55,082	69,244	0.79548		

d. Agreeableness

Variable	SumSq	DF	MeanSq	F	pValue
age	7.34	1	7.34	7.82	0.005
handedness	2.29	2	1.14	1.22	0.296
gender	962.66	2	481.33	512.85	<0.001
language	10.17	1	10.17	10.83	<0.001
ethnicity	45.56	1	45.56	48.54	<0.001
country	0.025	1	0.03	0.027	0.870
education	394.89	4	98.72	105.19	<0.001
income	250.49	11	22.77	24.26	<0.001
occupation	61.91	5	12.38	13.19	<0.001
Error	64,988	69,244	0.94		

e. Conscientiousness

Variable	SumSq	DF	MeanSq	F	pValue
age	31.90	1	31.90	33.41	<0.001
handedness	7.88	2	3.94	4.13	0.016
gender	211.97	2	105.99	111.01	<0.001
language	2.55	1	2.55	2.67	0.102
ethnicity	1.73	1	1.73	1.81	0.178
country	4.05	1	4.05	4.24	0.040
education	30.64	4	7.66	8.02	<0.001
income	244.71	11	22.25	23.30	<0.001
occupation	87.85	5	17.57	18.40	<0.001
Error	66,111	69,244	0.95		

occupation, income first language, country of residence, and ethnicity all had significant associations, albeit with varying patterns, across the five personality scores [28–31]. In addition, severity and prevalence of psychiatric and neurological conditions also covary with some of these factors [17,18,32,33].

3. Results

Overall demographic characteristics of the sample are shown in Table 2. Across the sample, in terms of report of prior diagnoses from a healthcare professional, 4854 participants reported depression but not anxiety, 4682 reported anxiety but not depression, 6047 reported both depression and anxiety, 368 reported ADHD, 476 reported OCD, 453 reported bipolar disorder, 2305 reported having other psychiatric condition(s). For neurological conditions (again report of prior diagnoses from a healthcare professional), 978 participants reported a learning disability, 298 reported multiple sclerosis, 495 reported having had a stroke, 159 reported Parkinson's disease, 124 reported having had a traumatic brain injury, and 2603 reported having another neurological condition.

3.1. Correspondence of full and abbreviated scales

We consider a large effect size as a robust and consistent relationship between variables, often indicated by correlation coefficients close to 1 or − 1. Medium effect sizes suggest a moderate relationship (correlation coefficients around 0.5), while small effect sizes indicate a weak association (coefficients around 0.2 or lower). Negligible effect sizes imply little to no relationship (coefficients close to 0) [34]. There were strong correlations (i.e. of extremely large effect size) between scores from the full-scale BFI and the current Abbreviated BFI (Table 3, Fig. 1): Openness ($r(29,896) = 0.95$, $p < 0.001$), Extraversion ($r(29,896) = 0.94$, $p < 0.001$), Neuroticism (emotional instability) ($r(29,896) = 0.91$, $p < 0.001$), Agreeableness ($r(29,896) = 0.94$, $p < 0.001$), and Conscientiousness ($r(29,896) = 0.86$, $p < 0.001$). By contrast, correlations between scores from the full-scale Big Five and previously recommended 10-item version [6] (supplementary material Table 2) were lower for some factors: Openness ($r(29,896) = 0.92$, $p < 0.001$), Extraversion ($r(29,896) = 0.90$, $p < 0.001$), Neuroticism (emotional instability) ($r(29,896) = 0.84$, $p < 0.001$), Agreeableness ($r(29,896) = 0.75$, $p < 0.001$), Conscientiousness ($r(29,896) = 0.72$, $p < 0.001$). The correlations between the full-scale Big Five and the 5-item version (supplementary material Table 3) were lower again: Openness ($r(29,896) = 0.86$, $p < 0.001$), Extraversion ($r(29,896) = 0.80$, $p < 0.001$), Neuroticism (emotional instability) ($r(29,896) = 0.71$, $p < 0.001$), Agreeableness ($r(29,896) = 0.76$, $p < 0.001$), Conscientiousness ($r(29,896) = 0.72$, $p < 0.001$).

3.2. Retest reliability

There were strong positive correlations between factor scores for the Abbreviated Big 5 at Time 1 and Time 2 for Openness ($r(10,370) = 0.84$, $p < 0.001$), Extraversion ($r(10,370) = 0.83$, $p < 0.001$), Neuroticism (emotional instability) ($r(10,370) = 0.81$, $p < 0.001$), Agreeableness ($r(10,370) = 0.81$, $p < 0.001$), and Conscientiousness ($r(10,370) = 0.73$, $p < 0.001$) (Table 4). Visualisation Bland–Altman plots [35] showed little evidence of bias in scores across timepoints (Fig. 2).

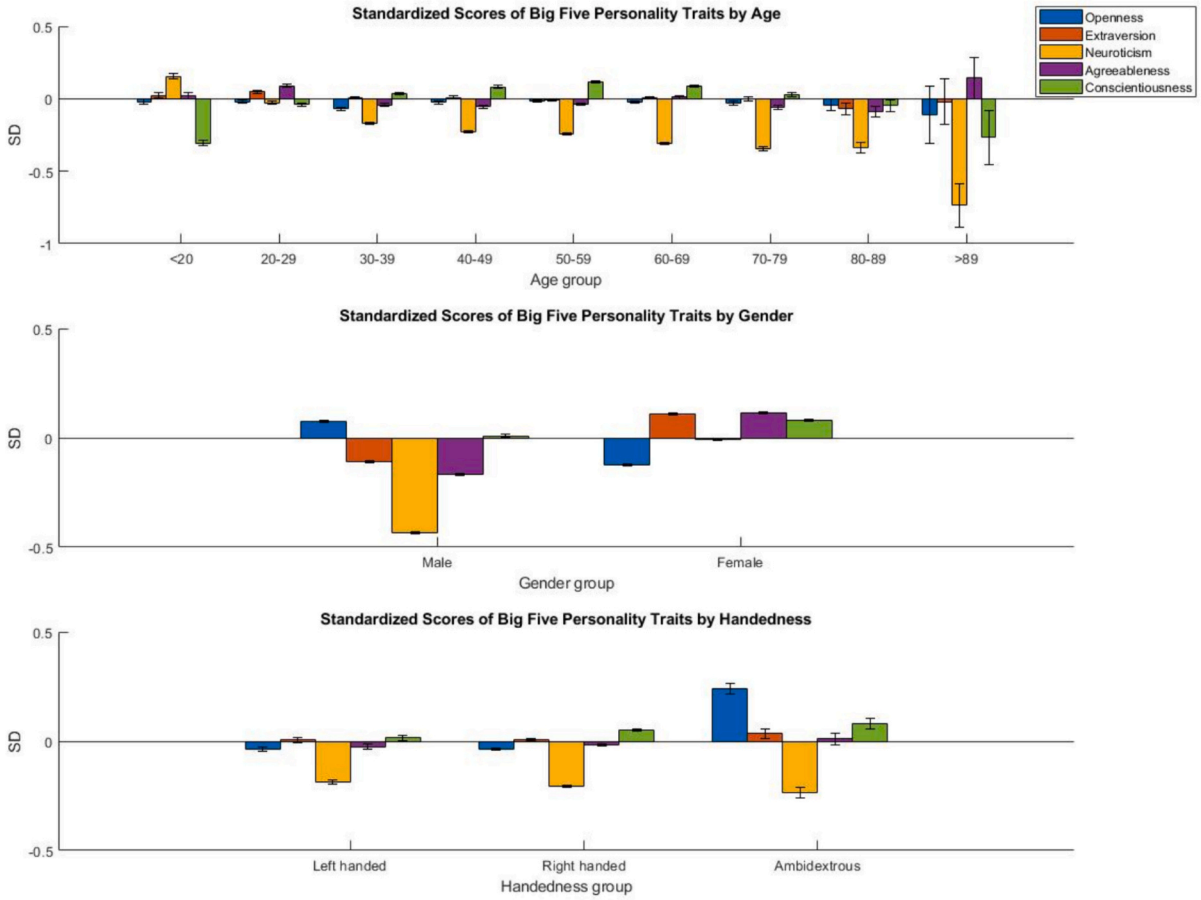


Fig. 3. Personality traits including Openness, Extraversion, Neuroticism (emotional instability), Agreeableness, Conscientiousness as a function of age, gender, and handedness. There were associations of age gender and handedness with personality traits but the association between age and Openness, Extraversion, and Agreeableness did not show a clear pattern. There also were relationships between gender and all dimensions of personality traits with males having significant higher scores in Openness but lower scores in Extraversion, Neuroticism (emotional instability), Agreeableness, and Conscientiousness in females. Finally, handedness was related to Openness and Conscientiousness but not Extraversion nor Agreeableness. Y axes = mean factor scores in SD units.

3.3. Normative modelling

Linear models were trained to statistically predict personality scores from demographic variables. Results are presented in full in Table 5. In brief, there were statistically significant associations of age with all five personality traits. Neuroticism (emotional instability) showed a robust decrease with age within the medium effect size range, while Conscientiousness presented a small inversed U-shaped relationship. Openness, Extraversion, and Agreeableness associations with age were of negligible scale (Fig. 3). There were also statistically significant associations between gender and all traits. These were mostly within the small to negligible effect size range, with males scoring higher in Conscientiousness and Openness, while females exhibited higher scores in Extraversion, Agreeableness and Neuroticism (emotional instability) (Fig. 3). Handedness had negligible scaled but statistically significantly associations with Neuroticism (emotional instability), and Conscientiousness. Ambidextrous individuals showed higher Openness scores (Fig. 3).

Language was statistically significantly related to Openness, Neuroticism (emotional instability), and Agreeableness, with English speakers having marginally higher Openness and lower Neuroticism (emotional instability) scores (Fig. 4). Country of residence was statistically significantly associated with Openness and Conscientiousness, where non-UK residents had marginally higher openness scores (Fig. 4). Educational attainment was significantly associated with all personality traits, with individuals holding a PhD scoring highest in Openness and Conscientiousness but lowest in Extraversion. Those with a university degree had the highest Agreeableness scores, and Neuroticism (emotional instability) scores were similar across all education levels (Fig. 4).

Occupation was linked with Openness, Extraversion, Neuroticism (emotional instability), and Conscientiousness. Unemployed, disabled, or sheltered individuals generally scored higher in these traits, with the exception of Neuroticism (emotional instability) and Conscientiousness where workers and retirees exhibited lower and higher scores, respectively (Fig. 5). Income was significantly associated with all traits, displaying positive correlations with Openness, Extraversion, and Conscientiousness and negative correlations with Neuroticism (emotional instability) and Agreeableness (Fig. 5). Ethnicity was significantly associated with Extraversion, Neuroticism (emotional instability), and Agreeableness, although differences between groups were of small-negligible effect size (Fig. 5).

3.4. Associations of the abbreviated Big 5 with psychiatric and neurological conditions

Applying *t*-tests to determine differences in the Abbreviated Big 5 of people reporting clinical conditions (Table 6, Fig. 6) indicated substantial but distinct differences in personality DfE scores for psychiatric groups relative to predictions of the normative model. We considered these effect sizes suggested by Cohen [34] (0.2 = small, 0.5 = medium, and 0.8 = large). Values smaller than 0.2 are considered as negligible.

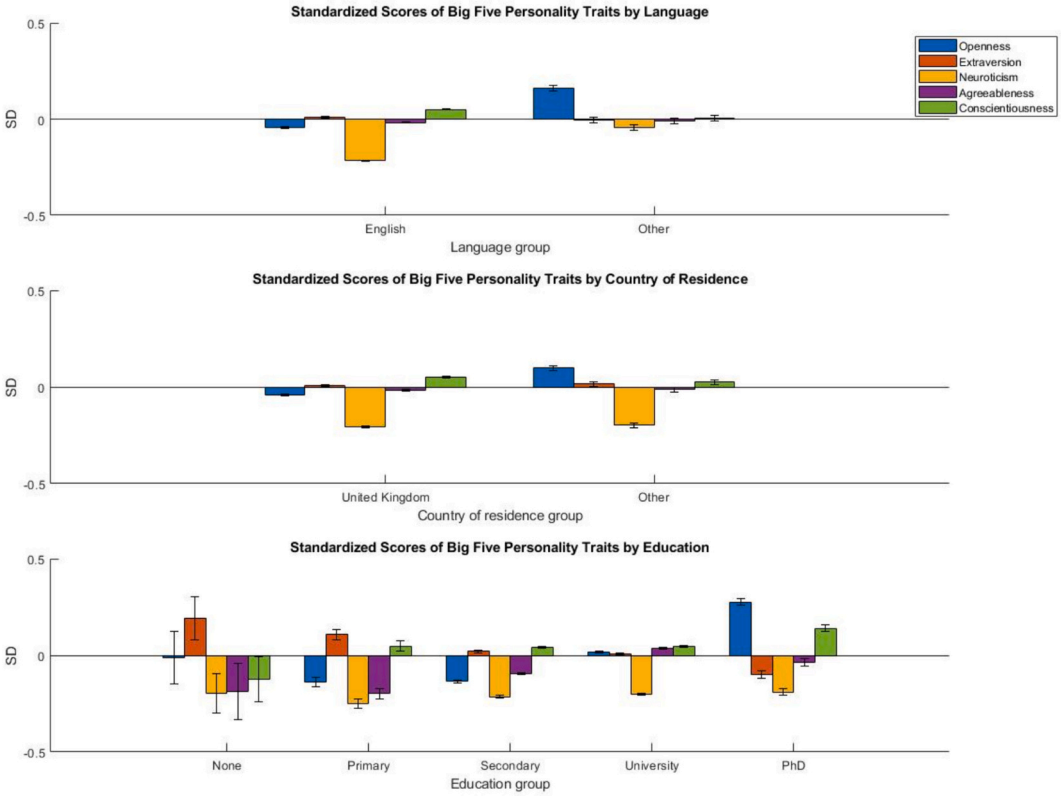


Fig. 4. Personality traits including Openness, Extraversion, Neuroticism (emotional instability), Agreeableness, Conscientiousness as a function of language, country of residence, and education. There were associations between language and Openness, Neuroticism (emotional instability), and Agreeableness but not Extraversion nor Conscientiousness. There were significant associations between country of residence and Openness and Conscientiousness but not on Extraversion, Neuroticism (emotional instability), and Agreeableness. Education was associated with all personality traits with people with PhD generally having higher Openness, Conscientiousness, and Agreeableness scores but a lower Extraversion score. Y axes = mean factor scores in SD units.

Neuroticism (emotional instability) had the strongest association, with increased mean scores in the medium to large effect size range for all psychiatric conditions. People with ADHD and bipolar disorder showed heightened scores in the medium effect size range for Openness, with smaller but statistically significant increases for depression, anxiety, co-existing depression and anxiety and OCD. People with ADHD showed a large effect size reduction in Conscientiousness scores, with depression, co-existing depression and anxiety, and bipolar disorder all showing small effect size reductions. Finally, people with ADHD showed heightened Extraversion scores in the medium effect size range. All other comparisons were either statistically non-significant or of negligible effect size.

Personality scores had variable and generally weaker relationships with neurological and developmental conditions (Table 6, Fig. 6). People with learning disabilities showed elevated Neuroticism (emotional instability) scores in the medium effect size range, with smaller elevation of scores for traumatic brain injury and multiple sclerosis. Openness also showed small elevations in score for learning disability and traumatic brain injury patients. People with Parkinson's disease showed medium sized reductions in Conscientiousness scores, with smaller reductions for people with learning disabilities. Finally, people with Parkinson's disease showed small-scaled reductions in Agreeableness. All other comparisons were either of negligible effect size or were statistically non-significant.

4. Discussion

These results demonstrate at large population scale that the 18-item version of the BFI is suitable for widespread use (including online deployment as part of a broader multi-scale survey), with superior psychometric validity relative to previously developed abbreviated scales and noteworthy patterns of association with psychiatric and neurological conditions.

The psychometric validity of the 18-item BFI was confirmed through robust correlations between component scores calculated from the full BFI and the shortened version, which were higher than for previously 10 item version [7] and the 5-item version selected based on the 5 most heavily loaded questions for each factor, especially for Openness, Conscientiousness and Agreeableness. They also were stronger than previously reported for 10-item [7], 15-item [11], and 20-item scales. Furthermore, the 18-item scale outperformed other shortened versions in retest reliability across two-time points than the 10-item [7] and 15-item [11] scales.

The observation of significantly scaled associations of demographic factors with personality trait scores highlights the importance of applying normative modelling when evaluating associations with other population factors such as psychiatric conditions. Consistent with many other studies [28–31], age, gender, education, occupation, income first language, country of residence, and ethnicity all had significant associations, albeit with varying patterns, across the five personality scores. This finding, which is consistent with previous studies examining socio-demographic factor associations with the full BFI, presents a challenge because the severity and prevalence of psychiatric and neurological conditions also covary with some of these factors [17,18,32,33]. Notably, the very large scale of data that can be collected efficiently and affordably online offers the potential to develop more detailed normative models that account for many such factors, or to use sub-sampling where individuals are selected from the large normative database that closely match patients according to those dimensions.

The large scale of the available data also enabled a transdiagnostic

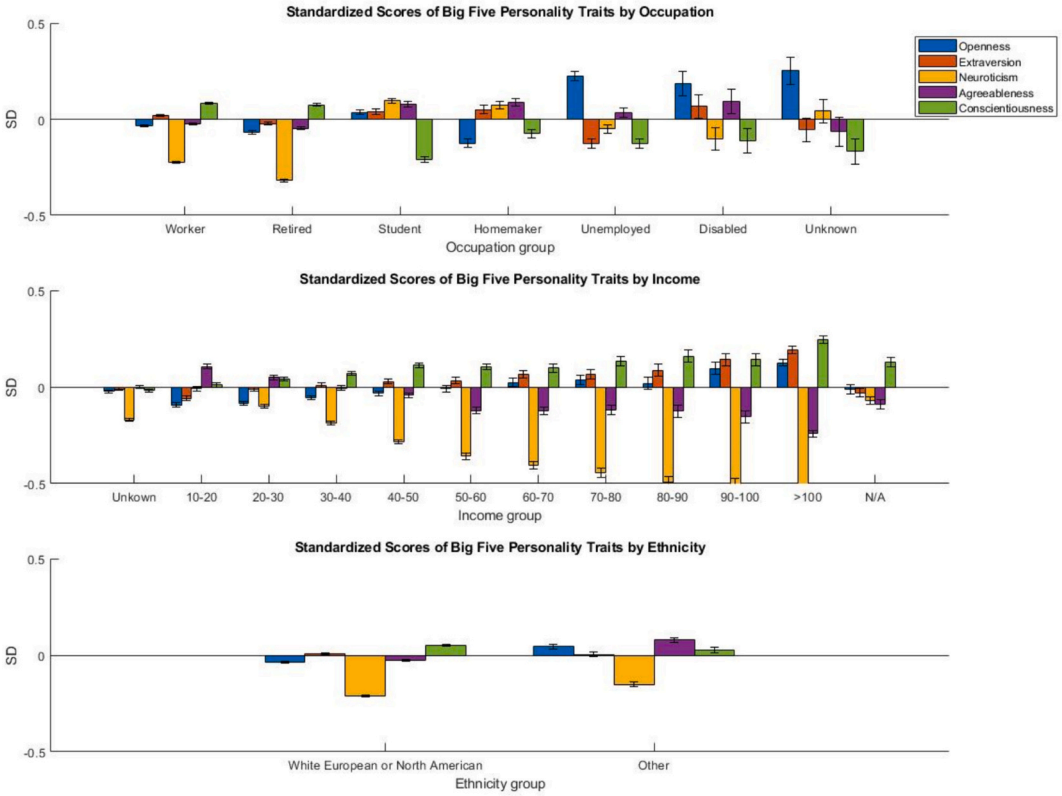


Fig. 5. Personality traits including Openness, Extraversion, Neuroticism (emotional instability), Agreeableness, Conscientiousness as a function of occupation, income, and ethnicity. There were associations at various directions between occupation groups and all personality traits. Income was also associated with all personality traits with Openness, Extraversion, and Conscientiousness positively related to income and Neuroticism (emotional instability) and Agreeableness negatively associated with income. Ethnicity is associated with Extraversion, Neuroticism (emotional instability), and Agreeableness, but not with Openness nor Conscientiousness. Y axes = mean factor scores in SD units.

analysis of personality trait associations with neurological and psychiatric conditions. After applying normative modelling, Neuroticism (emotional instability) had the strongest and most general profile, showing medium to strong scaled associations with all psychiatric conditions examined. This finding accords well with the broader literature on such associations, where Neuroticism (emotional instability) has been highlighted as having robust psychiatric associations [17–24]. Other personality traits had more nuanced relationships with psychiatric conditions. For example, higher Openness was most robustly associated with ADHD, bipolar disorder and OCD. Heightened Extraversion was strongly and specifically related to ADHD. Conscientiousness had a strong negative association with ADHD and more moderate associations with all three depression categories. Agreeableness had only small to negligible scaled associations with psychiatric disorders. These differential patterns of association may be partly explained by variation in brain structure and/or function related to specific traits. For instance, Neuroticism (emotional instability) has been associated with brain regions involved in threat, punishment, and negative affect, while Agreeableness is thought to be related to brain regions processing information associated with others. Conscientiousness has been linked to brain regions associated with planning and controlling behavior, whereas Extraversion has been associated with brain regions involved in reward information processing [36].

The patterns of association with psychiatric conditions align with some but not all such associations reported in the literature [17–24]. This may reflect that some past studies have analysed associations with symptoms [17,18] as opposed to participant-reported clinical labels, which were the focus here, though we also note the large population scale of the current study. This discrepancy highlights the challenge of understanding the causal relationships between personality and psychiatric conditions, a topic that remains under debate [13,15,16].

In accordance with past literature, personality traits in neurological groups showed generally weaker associations than psychiatric conditions and varied more according to condition. Learning disabilities appeared more similar to psychiatric conditions insofar as Neuroticism (emotional instability) was elevated to a significant degree along with Openness, whereas Conscientiousness was lower. In patients with traumatic brain injury, only Openness was significantly high. In patients with multiple sclerosis, although the Neuroticism (emotional instability) score was statistically high, which is consistent with some past research [25], the scale of the effect was very small. The observation of decreased Agreeableness and Conscientiousness in people with Parkinson's disease aligns with some past studies, though did not observe the expected lower levels of Openness and Extraversion and furthermore, effect sizes for stroke were of negligible scale [26,27].

Abbreviated scales of the type validated here are likely to be of particular use in clinical studies, and in longitudinal online studies, especially where multiple different scales are deployed, as shorter scales can reduce participant discouragement, fatigue, and inattention. Potential applications include longitudinal research with many time points, where online delivery of a brief scale alongside measures of mood or other psychological/neurological symptoms could enable causal relationships of personality traits with the development of psychiatric and neurological conditions or resilience to environmental stressors to be examined efficiently in fine temporal detail and at large population scale.

While our study provides a useful new short version of the BFI, it has some limitations that warrant consideration. First, the self-reported nature of participants' psychiatric and neurological conditions may introduce noise when evaluating association strengths; and of course is

Table 6
The results of the *t*-tests on the differences in Openness, Extraversion, Neuroticism (emotional instability), Agreeableness, and Conscientiousness between clinical groups and their expected scores in standard deviation units.

Group	Openness	Extraversion	Neuroticism (emotional instability)	Agreeableness	Conscientiousness
Depression	$t = 5.62$, $p < 0.001$, ES = 0.08	$t = -2.74$, $p < 0.01$, ES = -0.04	$t = 50.46$, $p < 0.001$, ES = 0.61	$t = -1.20$, $p < 0.05$, ES = -0.03	$t = 14.79$, $p < 0.001$, ES = -0.21
Anxiety	$t = 8.45$, $p < 0.001$, ES = 0.13	$t = -0.92$, $p = 0.36$, ES = -0.01	$t = 73.46$, $p < 0.001$, ES = 0.84	$t = 5.41$, $p < 0.001$, ES = 0.08	$t = -5.81$, $p < 0.001$, ES = -0.08
Depression + anxiety	$t = 10.97$, $p < 0.001$, ES = 0.15	$t = -10.31$, $p < 0.001$, ES = -0.13	$t = 102.35$, $p < 0.001$, ES = 1.05	$t = 4.26$, $p < 0.001$, ES = 0.06	$t = -15.31$, $p < 0.001$, ES = -0.20
ADHD	$t = 7.49$, $p < 0.001$, ES = 0.50	$t = 7.09$, $p < 0.001$, ES = 0.37	$t = 11.92$, $p < 0.001$, ES = 0.65	$t = 0.38$, $p = 0.70$, ES = 0.02	$t = -17.64$, $p < 0.001$, ES = -0.90
OCD	$t = 5.50$, $p < 0.001$, ES = 0.22	$t = 0.37$, $p = 0.71$, ES = 0.01	$t = 32.97$, $p < 0.001$, ES = 0.98	$t = 0.07$, $p = 0.94$, ES = 0.00	$t = -2.19$, $p < 0.05$, ES = -0.08
Bipolar	$t = 7.56$, $p < 0.001$, ES = 0.41	$t = 1.85$, $p = 0.06$, ES = 0.08	$t = 20.97$, $p < 0.001$, ES = 0.89	$t = 0.87$, $p = 0.39$, ES = 0.05	$t = -6.58$, $p < 0.001$, ES = -0.34
Other	$t = 11.82$, $p < 0.001$, ES = 0.28	T = -1.28, $p = 0.20$, ES = -0.03	$t = 53.12$, $p < 0.001$, ES = 0.99	$t = 1.39$, $p = 0.17$, ES = 0.03	$t = -9.68$, $p < 0.001$, ES = -0.22
Learning disability	$t = 9.42$, $p < 0.001$, ES = 0.32	$t = -2.17$, $p < 0.05$, ES = -0.07	$t = 12.68$, $p < 0.001$, ES = 0.42	$t = 2.37$, $p < 0.05$, ES = 0.09	$t = -8.46$, $p < 0.001$, ES = -0.28
Multiple sclerosis	$t = -0.42$, $p = 0.68$, ES = -0.03	$t = 0.31$, $p = 0.76$, ES = 0.02	$t = 2.08$, $p < 0.05$, ES = 0.12	$t = 0.24$, $p = 0.81$, ES = 0.02	$t = -1.39$, $p = 0.17$, ES = -0.09
Stroke	$t = 0.38$, $p = 0.70$, ES = 0.02	$t = 1.04$, $p = 0.30$, ES = 0.04	$t = -1.03$, $p = 0.30$, ES = -0.04	$t = -2.22$, $p < 0.03$, ES = -0.11	$t = -1.98$, $p < 0.05$, ES = -0.09
Parkinson's disease	$t = -1.72$, $p = 0.09$, ES = -0.17	$t = 0.48$, $p = 0.63$, ES = 0.04	$t = -0.19$, $p = 0.85$, ES = -0.01	$t = -3.71$, $p < 0.001$, ES = -0.38	$t = -6.01$, $p < 0.001$, ES = -0.48
Traumatic brain injury	$t = 2.96$, $p < 0.01$, ES = 0.29	$t = -1.22$, $p = 0.22$, ES = -0.10	$t = 1.75$, $p = 0.08$, ES = 0.17	T = -0.23, $p = 0.81$, ES = -0.02	$t = -1.89$, $p = 0.06$, ES = -0.15
Other	$t = 3.32$, $p < 0.001$, ES = 0.07	$t = -3.69$, $p < 0.001$, ES = -0.07	$t = 16.82$, $p < 0.001$, ES = 0.33	$t = -0.42$, $p = 0.66$, ES = -0.01	$t = -6.27$, $p < 0.001$, ES = -0.13

*p < 0.05, **p < 0.01, ***$p < 0.001$.

not as robust as conducting in-person clinical interviews. Second, the cross-sectional analysis did not seek to establish how psychiatric/neurological disorders relate to personality traits in terms of directions of influence/causality. This would require repeated data collection across a longer timescale, which we note is more feasible at scale with abbreviated online scales such as the one evaluated here. Our study focused on a specific UK sample and clinical populations, which may limit the generalisability of the findings to other populations and contexts. Third, the study measured lifetime mental and neurological conditions rather than being able to differentiate history of such diagnoses from current diagnoses. Future research should further validate the short Big Five scale in diverse cultures and clinical groups. Moreover, the current study chose its 18-item subset solely based on the strength of factor loadings. However, a significant drawback of this method is its tendency to result in a limited range of item content. For instance, the shortened Extraversion scale mainly focuses on talkativeness ("is talkative" vs. "tends to be quiet" and "is reserved"), lacking items that gauge assertiveness or enthusiasm. Likewise, the condensed Neuroticism (emotional instability) scale primarily measures anxiety ("worries a lot" vs. "is relaxed, handles stress well" and "is emotionally stable, not easily upset"), without addressing feelings of sadness or depressive affect. It is worth noticing that there are other possible item-selection methods with different advantages and disadvantages. Finally, it is worth mentioning that there are recently developed tools that may be better for measuring personality with enhanced bandwidth, fidelity, and predictive power [3,37].

In summary, this study evaluated and validated an 18 item BFI variant that can now be deployed across a range of future studies, including but not limited to online multiscale surveys. It showed sound psychometric properties and meaningful associations with a range of psychiatric and neurological conditions. The analyses of associations between personality traits and psychiatric/neurological conditions at large population scale also highlight the complex interplay between personality and mental health. The validated brief format offers an efficient means of collecting large-scale longitudinal data at minimal cost and may prove useful for research in large scale longitudinal studies, e.g., probing the causal relationships between personality traits and the development and progression of psychiatric disorders.

CRediT authorship contribution statement

Weixi Kang: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jeggan Tiego:** Writing – review & editing, Writing – original draft, Validation, Investigation, Conceptualization. **Peter J. Hellyer:** Software, Resources, Methodology, Data curation, Conceptualization. **William Trender:** Software, Resources, Methodology, Data curation, Conceptualization. **Jon E. Grant:** Resources, Methodology, Conceptualization. **Samuel R. Chamberlain:** Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Conceptualization. **Adam Hampshire:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation,

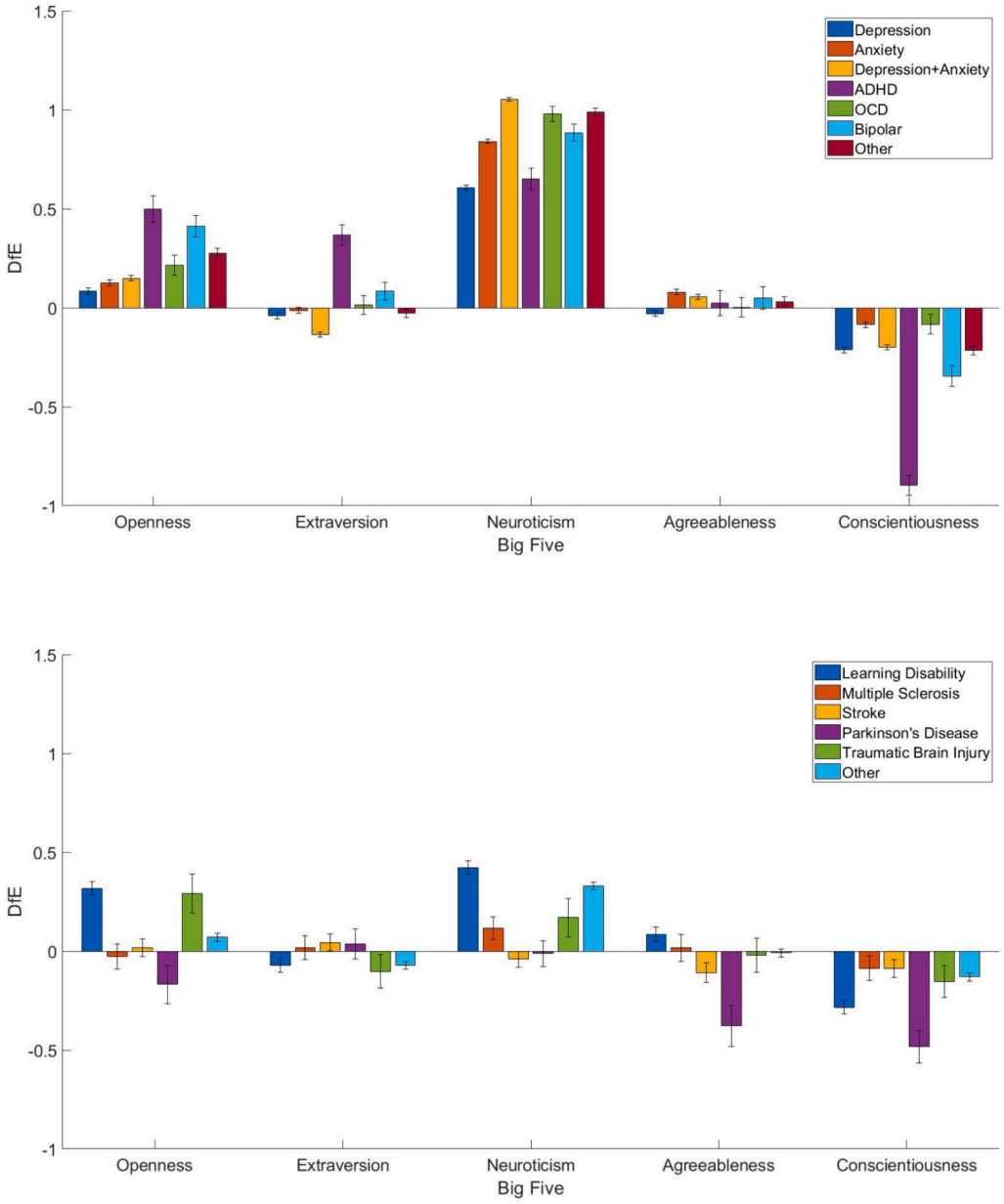


Fig. 6. The effect size of psychiatric and neurological conditions on personality. Size of associations between personality scores from the 18 item BFI with self-report previous formal diagnosis from a healthcare professional of psychiatric and neurological conditions after correcting for sociodemographic factors. Note strong associations between Neuroticism (emotional instability) and all psychiatric groups, and more selective associations of Openness and Conscientiousness scores with psychiatric conditions. Relationships between neurological conditions and personality traits were generally weaker than the relationships between psychiatric conditions and personality. Y axes = effect size in standard deviation units.

Conceptualization.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.comppsych.2024.152514.

References

[1] Goldberg LR. The structure of phenotypic personality traits. Am Psychol 1993;48(1):26.
[2] Soto CJ, Jackson JJ. Five-factor model of personality. Oxford University Press; 2013.
[3] John OP, Donahue EM, Kentle RL. The big five inventory - versions 4a and 54. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research; 1991.
[4] Laros Jacob Arie, de Souza Peres Alexandre José, de Andrade Josemberg Moura, Passos Maria Fabiana Damásio. Validity evidence of two short scales measuring the Big Five personality factors. Psicologia 2018;31.
[5] Caprara GV, Barbaranelli C, Livi S. Mapping personality dimensions in the Big Five model. Eur Rev Appl Psychol 1994;44(1):9–15.
[6] Gouveia VV, Araújo RDCR, de Oliveira ICV, Gonçalves MP, Milfont T, de Holanda Coelho GL, et al. A short version of the big five inventory (BFI-20): evidence on construct validity. Int J Psychol 2021;55(1):e1312.
[7] Rammstedt B, John OP. Measuring personality in one minute or less: a 10-item short version of the big five inventory in English and German. J Res Pers 2007;41(1):203–12.
[8] Saucier G. Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. J Pers Assess 1994;63(3):506–16.
[9] Kruyen PM, Emons WH, Sijtsma K. On the shortcomings of shortened tests: a literature review. Int J Test 2013;13(3):223–48.
[10] Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. Psychol Assess 2000;12(1):102.
[11] Hahn E, Gottschling J, Spinath FM. Short measurements of personality validity and reliability of the GSOEP big five inventory (BFI-S). J Res Pers 2012;46(3):355–9.

[12] Soto CJ, John OP. Short and extra-short forms of the big five inventory–2: the BFI-2-S and BFI-2-XS. J Res Pers 2017;68:69–81.
[13] Krueger RF, Tackett JL. Personality and psychopathology: working toward the bigger picture. J Personal Disord 2003;17(2) [Special issue), 109-128].
[14] Nigg JT. Temperament and developmental psychopathology. J Child Psychol Psychiatry 2006;47(3–4):395–422.
[15] Tackett JL. Evaluating models of the personality–psychopathology relationship in children and adolescents. Clin Psychol Rev 2006;26(5):584–99.
[16] Whittle S, Allen NB, Lubman DI, Yücel M. The neurobiological basis of temperament: towards a better understanding of psychopathology. Neurosci Biobehav Rev 2006;30(4):511–25.
[17] Hakulinen C, Elovainio M, Pulkki-Råback L, Virtanen M, Kivimäki M, Jokela M. Personality and depressive symptoms: individual participant meta-analysis of 10 cohort studies. Depress Anxiety 2015;32(7):461–70.
[18] Karsten J, Penninx BW, Riese H, Ormel J, Nolen WA, Hartman CA. The state effect of depressive and anxiety disorders on big five personality traits. J Psychiatr Res 2012;46(5):644–50.
[19] Perroud N, Hasler R, Golay N, Zimmermann J, Prada P, Nicastro R, et al. Personality profiles in adults with attention deficit hyperactivity disorder (ADHD). BMC Psychiatry 2016;16(1):1–9.
[20] Fardin MA, Nooripour R, Shirazi M, Farnam A, Arab A. Role of big five personality traits in obsessive-compulsive disorder and sleep quality among students. J Res Health 2017;7(6):1086–93.
[21] Barnett JH, Huang J, Perlis RH, Young MM, Rosenbaum JF, Nierenberg AA, et al. Personality and bipolar disorder: dissecting state and trait associations between mood and personality. Psychol Med 2011;41(8):1593–604.
[22] Jylhä P, Mantere O, Melartin T, Suominen K, Vuorilehto M, Arvilommi P, et al. Differences in neuroticism and extraversion between patients with bipolar I or II and general population subjects or major depressive disorder patients. J Affect Disord 2010;125(1–3):42–52.
[23] Nowakowska C, Strong CM, Santosa CM, Wang PW, Ketter TA. Temperamental commonalities and differences in euthymic mood disorder patients, creative controls, and healthy controls. J Affect Disord 2005;85(1–2):207–15.
[24] Sparding T, Pålsson E, Joas E, Hansen S, Landén M. Personality traits in bipolar disorder and influence on outcome. BMC Psychiatry 2017;17(1):1–10.
[25] Abedini SM, Montazeri S, Ghorban Shiroudi S. Comparing the big five personality factor in patients with multiple sclerosis and healthy individuals. J Mazandaran Univ Med Sci 2012;22(88):35–9.
[26] Dwan T, Ownsworth T. The Big Five personality factors and psychological well-being following stroke: a systematic review. Disabil Rehabil 2019;41(10):1119 1130.
[27] Santangelo G, Garramone F, Baiano C, D'Iorio A, Piscopo F, Raimo S, et al. Personality and Parkinson's disease: a meta-analysis. Parkinsonism Relat Disord 2018;49:67–74.
[28] Dahmann SC, Anger S. The impact of education on personality: Evidence from a German high school reform. 2014.
[29] Denissen JJ, Bleidorn W, Hennecke M, Luhmann M, Orth U, Specht J, et al. Uncovering the power of personality to shape income. Psychol Sci 2018;29(1): 3–13.
[30] Mann FD, DeYoung CG, Krueger RF. Patterns of cumulative continuity and maturity in personality and well-being: evidence from a large longitudinal sample of adults. Personal Individ Differ 2021;169:109737.
[31] Stake JE, Eisele H. Gender and personality. In: Handbook of gender research in psychology. New York, NY: Springer; 2010. p. 19–40.
[32] Ansseau M, Fischler B, Dierick M, Albert A, Leyman S, Mignon A. Socioeconomic correlates of generalized anxiety disorder and major depression in primary care: the GADIS II study (Generalized Anxiety and Depression Impact Survey II). Depress Anxiety 2008;25(6):506–13.
[33] Feigin VL, Vos T, Nichols E, Owolabi MO, Carroll WM, Dichgans M, et al. The global burden of neurological disorders: translating evidence into policy. Lancet Neurol 2020;19(3):255–65.
[34] Cohen J. Statistical power analysis for the behavioral sciences. Academic press; 2013.
[35] Giavarina D. Understanding bland altman analysis. Biochem Med 2015;25(2): 141–51.
[36] DeYoung CG, Hirsh JB, Shane MS, Papademetris X, Rajeevan N, Gray JR. Testing predictions from personality neuroscience: brain structure and the big five. Psychol Sci 2010;21(6):820–8.
[37] Soto CJ, John OP. The next big five inventory (BFI-2): developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. J Pers Soc Psychol 2017;113(1):117.