ELSEVIER

A spectral clustering approach to the structure of personality: Contrasting the FFM and HEXACO models

CrossMark

Sean Brocklebank [a,*], Scott Pauls [b], Daniel Rockmore [b,c], Timothy C. Bates [d,*]

[a] School of Economics, University of Edinburgh, United Kingdom
[b] Department of Mathematics, Dartmouth College, Hanover, NH, USA
[c] The Santa Fe Institute, Santa Fe, NM, USA
[d] Department of Psychology, University of Edinburgh, United Kingdom

ABSTRACT

Alternative analytic methods may help resolve the dimensionality of personality and the content of those dimensions. Here we tested the structure of personality using spectral clustering and conventional factor analysis. Study 1 analysed responses from 20,993 subjects taking the 300-item IPIP NEO personality questionnaire. For factor analysis, a five-factor solution recovered the FFM domains while the six-factor solution yielded only a small and hard to interpret sixth factor. By contrast, spectral clustering analysis yielded six-cluster solutions congruent with the HEXACO model. Study 2 analysed data from 1128 subjects taking the 100-item HEXACO-PI-R. Unambiguous support was found for a six-cluster solution. The psychological content of the 6 clusters and their relationship to the FFM domains is discussed.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Taxonomy is basic for any science – often being referred as the "facts" of a field, for which theories then compete to account (McCrae & John, 1992). The end of the 20th century saw the emergence of a broad consensus regarding the basic dimensions of personality as consisting of five orthogonal, broad bandwidth domains, with numerous facets clustered beneath these core domains. Variants such as the Five Factor Model (FFM) (McCrae & Costa, 1997, 2003) or the Big Five (Goldberg, 1990) have often been seen as opportunities to refine and redefine the orientation of domains within this space, but both support five basic domains of personality. Considering the question of whether there are additional basic factors of personality not already included among the FFM, McCrae and John (1992, p. 190) concluded that this "*appears increasingly unlikely, given the wealth of data in support of the comprehensiveness of the FFM*". However while the item space of large FFM inventories may be comprehensive, there remains the possibility that the contents of personality may usefully be arranged in terms of other than five basic domains.

One actively researched alternative to the FFM – the HEXACO model – suggests that personality consists of six rather than five basic domains, with the additional domain being one of Honesty–Humility (Ashton & Lee, 2007). Models with fewer domains, for instance Eysenck's PEN structure, retain some support also (e.g. Tilopoulos, Pallier, & Coxon, 2010). A second area of research focuses on psychological content: the characteristics of high and low poles on each of the basic domains of personality (e.g. Costa & McCrae, 1998). Alternative methods to analyse the structure contained in personality data may yield differential support or elucidate differences between these competing models.

Measuring the structure of personality requires both that the questionnaire or other data adequately sample the universe of personality variation, and that the analytic methods applied to these data are sensitive to this information. The Five Factor Model emerged from a research programme designed, above all, to generate a comprehensive sampling of human personality (McCrae & John, 1992). If the FFM items universe is comprehensive, then adequate analytic tools should then reveal the basic dimensions of personality, even if these are fewer or larger than five. In Studies 1 and 2, we apply one such tool – spectral clustering – to two large (n = 20,993, n = 1126) data sets, one sampling the FFM domains, and one explicitly designed to sample the HEXACO domains.

Before proceeding to these two empirical studies using spectral clustering, we first outline the logic of this method, endeavouring to leverage the reader's existing knowledge of factor analysis.

* Corresponding authors.
   E-mail addresses: sean.brocklebank@ed.ac.uk (S. Brocklebank), tim.bates@ed.ac.uk (T.C. Bates).

1.1. Spectral clustering

Factor analysis and confirmatory factor analysis (Joreskog, 1969) remain the primary techniques for exploring structure in questionnaire data and are the foundation for the FFM. Of course a set of well known methods have emerged to selecting number of factors in exploratory factor analysis both visually (Cattell, 1966) and analytically (Horn, 1965; Revelle & Rocklin, 1979). A vast range of research has also focussed on contrasting distinct solutions for personality structure via exploratory and confirmatory techniques (e.g. Church & Burke, 1994), and determining criteria for evaluating the structure of personality. Confirmatory factor analytic studies of the FFM have suggested poor fit to the theorised model (Gignac, Bates, & Jang, 2007). Newer exploratory SEM approaches, however, which relax the strict criteria of CFA, allowing an EFA-structure at the item level, indicate that well-fitting five-factor models can be constructed within this framework (Marsh et al., 2010). Interest in exploratory SEM is growing, as is interest in determining criteria for evaluating structure more generally (Hopwood & Donnellan, 2010). Alongside conventional factor analysis-based approaches, alternative analytic strategies widely used outside personality, been used to address the question of the basic structure of personality (Tiliopoulos et al., 2010). In particular, spectral clustering (used in the present report) has emerged as a valuable tool among clustering techniques (Von Luxburg, 2007).

While spectral clustering shares with factor analysis the basic objective of creating a low-dimensional representation of data, it differs substantially in its optimisation target (see e.g., Braun, Leibon, Pauls, & Rockmore, 2011; Leibon, Pauls, Rockmore, & Savell, 2008; Ng, Jordan, & Weiss, 2001). Factor analysis operates on the correlation matrix and tries to replicate it as closely as possible in a smaller number of dimensions (Cattell, 1978; Spearman, 1927). Spectral clustering also has the objective of summarising the data in fewer dimensions, but differs from factor analysis both because it operates on a spatial transformation of the correlation matrix (where each item becomes a point in space) and because it attempts to segregate these items into discrete clusters so that the similar items are kept in the same cluster, while dissimilar items are put into different clusters (see e.g., Braun et al., 2011; Leibon et al., 2008; Ng et al., 2001).

Aspects of spectral clustering have close analogues with factor analysis. For instance, spectral clustering has a cluster-number parameter (k) corresponding to the number of factors requested for extraction in conventional factor analysis. Other aspects, however, are quite distinct. In particular, spectral clustering translates correlations among items into distances (sometimes described as adjacencies – the inverse of distances), and can transform these distances to emphasise particular kinds of relationships by varying what is known as the scale parameter (sigma).

The explanation below is designed to give the casual reader a good intuition about spectral clustering; in addition, online Supplementary Material is provided which contains a technical description, a full list of references and a runnable Matlab implementation of the procedures for those wishing to implement the procedure and/or to learn more about it. Readers interested in the formal mathematical detail should of course consult the source references.

2. Mapping correlations as points in space

The algorithm underlying spectral clustering converts the correlation matrix into a spatial representation – points in space (see Fig. 1). Doing this requires taking the correlation matrix and converting each correlation into a measure of distance (i.e. how far apart is each pair of items?). The measure of distance has two important properties: (i) the measure cannot fall below zero (distances cannot be negative), and (ii) distances are smallest when correlations are largest (so that the items are close together in space), larger for independent items, and largest for negatively linked items.

The conversion of data from a conventional n × n correlation matrix of variables into spatial distances between these variables is shown in Fig. 1. As shown, items that correlate strongly positively are placed near to each other. Items having a strong negative correlation with each other are placed most far apart. Intermediate correlations translate into intermediate distances. The 3-item case depicted in Fig. 1 is chosen so that each item distance can be realised in the 2-D plane of the page. In the general case, mapping n items can require up to $n - 1$ spatial dimensions. With the translation from correlations into distances achieved, Fig. 2 shows how a spatial representation of items is split into clusters.

The clustering operation (see Fig. 2) creates solutions with k-clusters of items by cutting the paths connecting item-pairs. This generates sets of items each connected to each other within a cluster, but not connected to any item outside the cluster. This is an iterative procedure, which can take a considerable time, as multiple alternative cutting solutions are generated and compared against a criterion, namely to minimise the total length of remaining connections while forming k-clusters of items. This is shown in Fig. 2: It can be seen that this criterion results in the creation of clusters consisting of items that are close to one another, but distant from items in other clusters. We next discuss a feature of spectral clustering which allows it to re-scale the translation of correlations into distances, and which plays an important role in allowing this method to detect structure in data.

3. Scale parameter: sigma

A specific advantage of the spectral clustering algorithm is its inclusion of a scale parameter – sigma (see Fig. 3). By adjusting this parameter, it is possible to vary the relative weight placed on the weakest versus the strongest correlations when performing the optimisation. The scale parameter has some analogues in processes widely used to re-weight correlation matrices in other fields (Sammon, 1969), but has not been deployed in Factor Analysis of personality. Fig. 3 shows this mapping for two example values of sigma. The figure note details how sigma alters the translation of correlation into spatial distance for the example values of sigma.

The ability to reweight correlations is valuable for discovering and understanding structure within data. In particular, setting the scale parameter to a low value emphasises strong correlations among pairs of items. For personality data, the relevance of scale is particularly apparent if a questionnaire is thought to contain a small number of items strongly targeting a domain. For instance, in the present case, we expect a questionnaire designed primarily to assess the FFM domains to contain a relatively small number of Honesty–Humility items. If the HEXACO model is correct these items will nevertheless show strong correlations with each other, and relatively weak correlations with items in the FFM domains. Fig. 4 shows how a low value for sigma can correctly identify clusters represented by only a sparse set of valid items. The low values of sigma increase all inter-item distances but magnify large distances disproportionately more. This has the effect of increasing the distance between valid clusters so they can be more readily identified (this effect of sigma is material for Study 1 below, where we examine personality structure in a questionnaire believed to have sparse coverage of one domain).

When all domains are well represented in the original set of items, changing sigma will not necessarily have any effect on the
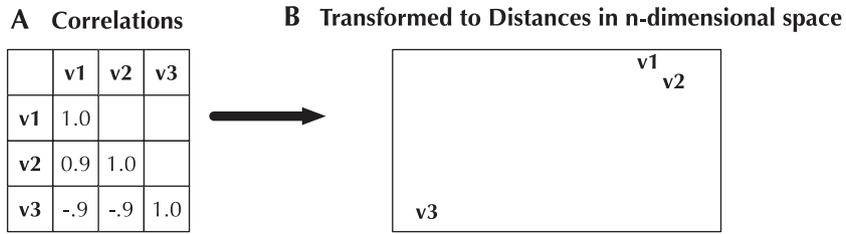
A   **Correlations**   **B**   **Transformed to Distances in n-dimensional space**

	v1	v2	v3
v1	1.0		
v2	0.9	1.0	
v3	-.9	-.9	1.0

Fig. 1. Mapping from correlation matrix input to *n*-dimensional distance. *Note:* Panel A shows a conventional input matrix, containing the *n* × *n* correlation matrix of variables, in this case v1–v3 (*n* = 3). Spectral analysis begins by translating these correlations into distances in an *n*-dimensional space. Items that correlate strongly positively are placed adjacent to one another. Items with strong negative correlations are placed distant from one another with intermediate correlations falling farther apart, and large negative relations most distant.
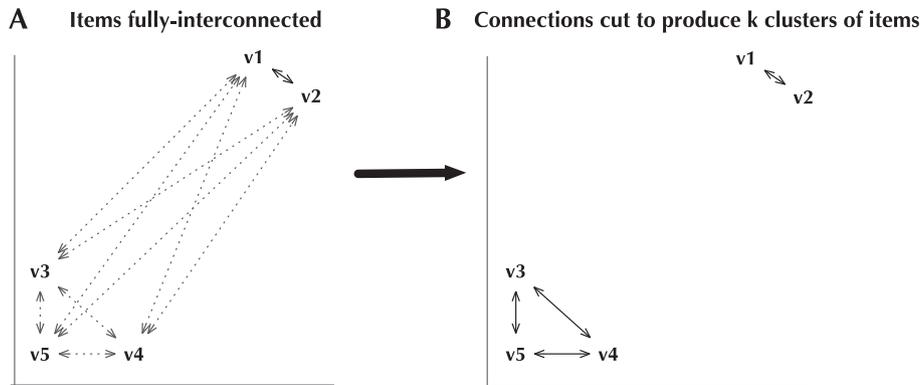
A   **Items fully-interconnected**   **B**   **Connections cut to produce k clusters of items**

Fig. 2. Cluster formation by cutting to form *k*-clusters of inter-connected items. *Note:* Clustering operates by iteratively seeking solutions with *k*-clusters of items by "cutting" paths between items. The clustering criterion is to minimise the total length of remaining connections (i.e., cutting the most distant relations) while forming *k*-clusters of items.
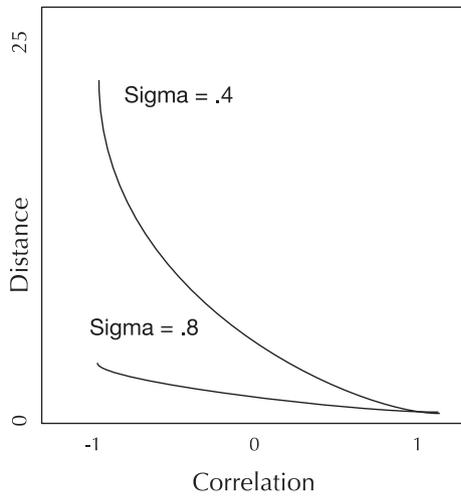
Fig. 3. Mapping correlations to distances at different values of the scale parameter. *Note:* Spectral clustering works on item distances in an *n*-dimensional space. Correlations (ranging from −1 to 1 as shown on the *x*-axis), are mapped such that high positive correlations map to short distances, and high negative correlations to the greatest distance. The sigma or scale parameter of spectral clustering allows a transform that differentially weights correlations according to their magnitude. At small values of sigma, the range of distance values is larger, thus increasing the importance of the strongest correlations (the weaker ones are effectively ignored). In summary: decreasing sigma increases the relative weight placed on the very strongest connections.

best-fitting value for cluster number (*k*) (see Fig. 5). With adequate numbers of items measuring each domain, it is possible to readily identify the corresponding cluster at all value of sigma (this effect is material for our hypotheses in Study 2 below).

Thus spectral clustering affords the possibility, but not the necessity, of discovering additional meaningful clusters. It can provide both a robustness check when compared to factor analysis, and, because of the control it offers over sensitivity to item relationships, can detect clusters sampled by relatively few items. Next, we apply this method to a large data set of IPIP NEO items.

4. Study 1

4.1. Introduction

In Study 1 we conducted factor analytic and spectral clustering analyses of responses to a battery of 300 items selected to represent the dimensions of the Five Factor Model (Johnson, 2005). This allows us to contrast support for the major competing personality models: FFM models, six-domain HEXACO models, and models arguing for other numbers of basic domains, such as Eysenck's three-factor PEN model (Eysenck, Eysenck, & Barrett, 1985). To the best of our knowledge, this is the first application of spectral clustering to personality data.

If five domains correctly describe personality, then we expect spectral clustering analysis to recover this, as factor analysis does. If, however, the six-domain HEXACO model (Ashton & Lee, 2005) is correct, and if the items of the NEO richly sample this domain, then spectral clustering should recover this six-cluster solution at all values of sigma, the scaling factor. Finally, if six clusters correctly characterise personality structure, but the IPIP NEO items only sparsely sample the Honesty–Humility domain, then the six cluster solution should fit well at lower values of the scaling parameter. This outcome emerges because small values of sigma optimise the detection of clusters that contain relatively few items sampling a domain, when these few items are weakly related to the main
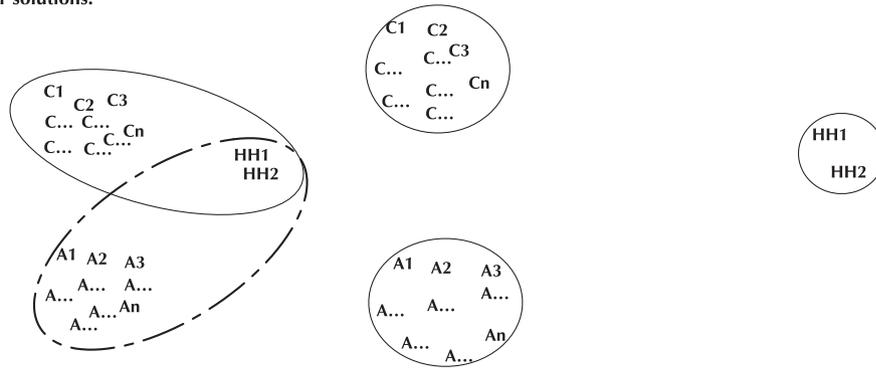
Fig. 4. Clustering of sparsely sampled domain with high- (panel A) versus low-values of sigma (panel B). *Note:* In this example, three clusters exist in nature. Two are richly represented by items in the data, and one (HH in this case) is only sparsely sampled. Panel A shows unstable 2-cluster solutions emerging in such data, with items from the weakly-sampled cluster forced into one or other of the two strongly-represented clusters. Panel B shows how small values of sigma, by moving items apart, reliably reveal the third cluster.
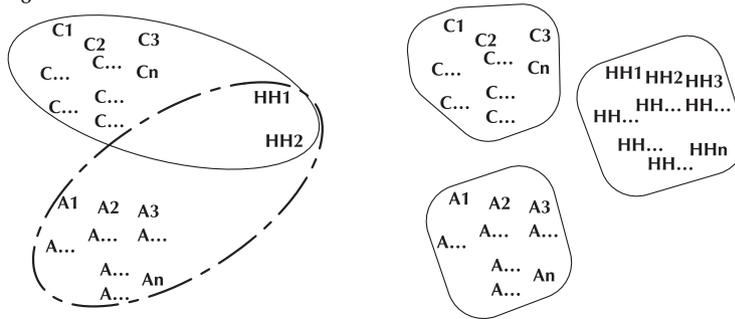

Fig. 5. Comparison of unstable cluster acquisition of a cluster sampled by few items (panel A), versus reliable acquisition of correct clusters at all values of sigma when data include sufficient markers for all clusters (panel B).

bulk of the items. In a questionnaire optimised to sample the domains of the five-factor model it is likely that relatively few questions exist which primarily sample Honesty–Humility, and thus this last outcome is the one we deemed most likely.

4.2. Method

Participants. For this analysis, we used data originally collected by John A. Johnson (2005). Johnson's study included over twenty thousand online responses to a five-factor personality inventory. Subjects for this study were not actively recruited – they either discovered the web site on their own or heard about it via word-of-mouth. In total, 23,994 subjects participated between August 6, 1999 and March 18, 2000. "Reported ages ranged from 10 to 99, with a mean age of 26.2 and SD of 10.8 years." (Johnson, 2005, p113). One of Johnson's aims was to discover effective strategies for detecting and excluding invalid data in online data acquisition. In the end, 20,993 of the original 23,994 participants' submissions were retained after excluding subjects for long strings of identical or missing responses and duplicate submissions detected via a algorithm comparing duplicate IP addresses, time intervals, and nickname overlap. The final sample consisted of 13,249 females (mean age 26.1 years) and 7,744 males (mean age 26.2 years).

Measures. Subjects completed 300 items from the International Personality Item Pool's (IPIP) NEO questionnaire, which has been

developed to measure the same constructs as the NEO-PI-R (Goldberg et al., 2006). The test includes five domain-level constructs: Neuroticism (N), Extraversion (E), Openness (O), Agreeableness (A), and Conscientiousness (C). Both the IPIP NEO and the NEO-PI-R measure six facets per domain. The mean correlation of facets on the IPIP proxy with corresponding facets in the NEO-PI-R is 0.94 after correcting for unreliability (Goldberg, 1999).

4.3. Analyses

Prior to analysis, scores on reverse-scored items were re-coded. Spectral clustering was done across a range of cluster numbers (k: 2:10) and across a range of values of sigma (from .4 to 1 in steps of .05). We did not consider values of sigma below 0.4 because below this threshold, the resulting network of items and their relations becomes disconnected (many of the distances go towards infinity). Best-fitting solutions were chosen according to cluster consistency evaluation. This method of evaluating fit involves generating solutions for all studied levels of k and sigma in the full data set. Additional solutions are then generated based on random samples of half the items. This is repeated 1000 times for each possible combination of k and sigma. The proportion of sampled items whose cluster classification differs from that in the clustering based on the full set of items indicates inconsistency. The best-fit value for k and for sigma is that which generates solutions that, across the 1000 repetitions, show the lowest proportion of

reclassified items. The analysis took several days of runtime on the computer.

4.4. Results

The goodness of fit measure used for parameter selection is the average level of item-misclassification in the resampling analyses. Fig. 6 shows the minimum percentage classification error across 100 runs for each value of sigma and k (the number of clusters, analogous to the number of factors in factor analysis). Darker shading indicates greater misclassification (bad fit), with lighter shading indicating less misclassification (good fit). The minimum error in each row is marked with a star; more than one cell may be starred in a row if the minima are not significantly different from one another. Table 1 gives numerical values of the misclassification rate and the standard error for various values of sigma and k for analyses of the IPIP data.

In spectral clustering, the value(s) of k that minimise misclassification error may vary depending on sigma. For the IPIP data, for all values of sigma greater than 0.7, the optimal cluster number (k) was five. However, for values of sigma smaller than 0.7, multiple values of k were not statistically significantly different from each other. For sigma between 0.4 and 0.7 minima emerged for both five and six cluster solutions (with the exception of sigma = 0.45 where five was not a minima). A three-cluster solution emerged in only one instance; likewise, a seven-cluster solution emerged as minima in a single instance.

Thus, while five-cluster solutions were preferred at high values of sigma, in line with the hypothesised sparse inclusion of valid Honesty–Humility items in the NEO IPIP inventory, at low values of sigma which emphasise these sparse but strong item clusters, six-cluster solutions fit the data as well or better.

4.5. Comparison of domains from spectral clustering and factor analysis

For the purposes of comparison with the results from clustering, exploratory factor analyses extracting five and six factors were

Table 1
Classification error data from Study 1, showing mean misclassification errors (and SE) as a function of sigma for each of the 3- to 7-cluster solutions.

Sigma	Classification error	Number of clusters				
		$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
.40	Misclassification rate	0.166	0.137	0.122	0.123	0.123
	Standard error	0.003	0.002	0.002	0.002	0.001
.45	Misclassification rate	0.127	0.15	0.133	0.122	0.142
	Standard error	0.003	0.003	0.002	0.002	0.002
.50	Misclassification rate	0.143	0.143	0.12	0.127	0.146
	Standard error	0.003	0.003	0.002	0.002	0.002
.55	Misclassification rate	0.169	0.14	0.13	0.125	0.142
	Standard error	0.003	0.003	0.002	0.002	0.002

Note: This table shows the mean and standard error of the proportion of misclassified items at each value of sigma and k. Each cell is based on 1000 simulations.

conducted, using a varimax rotation. In both cases, a factor classification was calculated for each item by assigning it to the factor on which the item has its highest loading following varimax rotation. This allowed a comparison of item-classifications between the factor analytic and spectral clustering solutions.

For the five-cluster and five-factor models, item classifications were very similar: 244 of the 300 items were assigned identically in both the factor and cluster solutions. Both solutions closely resembled the conventional Five Factor Model domains of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. By contrast, the six-cluster and six-factor solutions differed sharply. For the factor analytic solution, just seven (of a possible 300) items had their highest loading on the sixth factor. Moreover these were drawn from disparate facets, not related to theoretical models: (four items were initially located in the Self-Consciousness facet (N4) of Neuroticism, and one item from each of three additional facets of Agreeableness or Neuroticism). The six-cluster solution (see Table 2) yielded a substantial 38-item sixth cluster, with a coherent pattern of item sources: 12 items originally from the NEO-IPIP Conscientiousness domain, 20 items from Agreeableness, and six items from Neuroticism. Thus while the sixth factor was a small nuisance factor with a majority
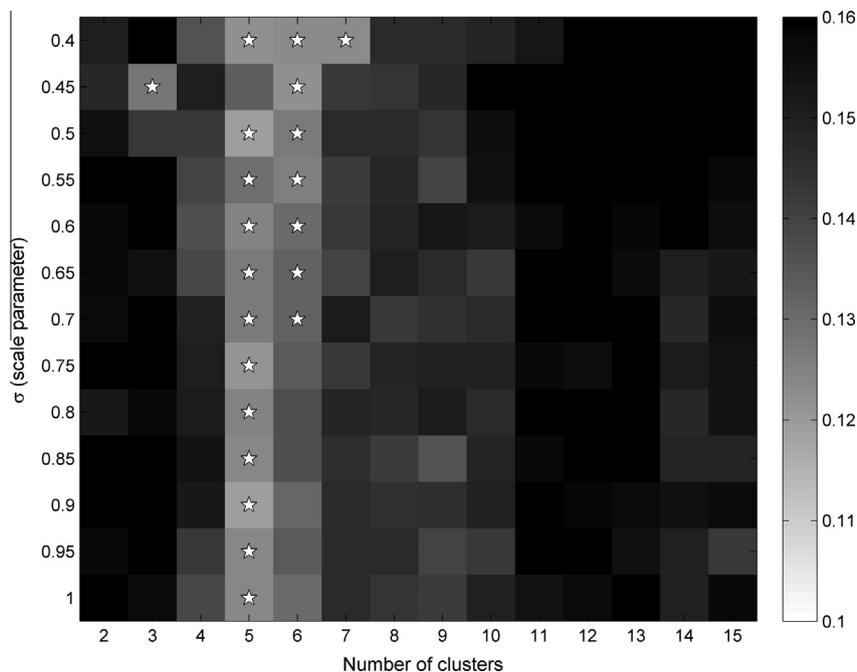

Fig. 6. Study 1: Mean of the minimum classification error percentages for all the trials for a given pair of k (the number of clusters) and σ (sigma—the scale parameter). *Note:* Stars indicate the lowest value in a given row (multiple stars are used when the values are not significantly different from one another).

Table 2
Six Cluster results: Numbers correspond to the number of items from each facet which were clustered in the corresponding 6-cluster domain (results computed at sigma = .5).

NEO IPIP facet		HEXACO domain					
Facet	Code	H–H	N	E	A	C	O
Anxiety	N1		9	1			
Anger	N2		10				
Depression	N3		9			1	
Self-Consciousness	N4		2	6		2	
Immoderation	N5	6	2			2	
Vulnerability	N6		8			1	1
Friendliness	E1			8	2		
Gregariousness	E2			10			
Assertiveness	E3			6		4	
Activity Level	E4					10	
Excitement-Seeking	E5			10			
Cheerfulness	E6		1	5	4		
Imagination	O1						10
Artistic Interests	O2			1	2		7
Emotionality	O3				5		5
Adventurousness	O4			1			9
Intellect	O5						10
Liberalism	O6						10
Trust	A1		1		9		
Morality	A2	5			5		
Altruism	A3				10		
Cooperation	A4	5	1		4		
Modesty	A5	10					
Sympathy	A6				10		
Self-Efficacy	C1		1			9	
Orderliness	C2					10	
Dutifulness	C3	4			1	5	
Achievement-Striving	C4					10	
Self-Discipline	C5					10	
Cautiousness	C6	8				2	

of members from a single facet, the sixth cluster was both substantial and meaningful. This was the case across multiple levels of sigma, with all six-cluster solutions from Study 1 being similar in item content, number of items contained, and resemblance to the HEXACO model. Table 2 (computed at sigma = 0.5) is representative of these.

In terms of psychological content, the majority (15 of 20) of Agreeableness items associated with the Honesty–Humility cluster were drawn from the A2 (Morality) and A5 (Modesty) facets: Example items assigned to the sixth cluster included "*Use flattery to get ahead*" (A2; reversed), "*Use others for my own ends*" (A2; reversed), "*Dislike being the centre of attention*" (A5), and "*Seldom toot my own horn*" (A5). Of the 12 Conscientiousness items that were reassigned to the sixth cluster, eight came from C6 (Cautiousness), including items such as "*Choose my words with care,*" and "*Make rash decisions*". The other four came from C3 (Dutifulness) including "*Break rules,*" and "*Get others to do my duties.*" The Neuroticism items clustered with Honesty–Humility all came from N5 (Immoderation), including items such as "*Rarely overindulge,*" and "*Never splurge.*" These further reinforce the similarity of the sixth cluster to Honesty–Humility.

The O3 facet of Openness split across two clusters. Seven items centred on compassion and empathy moved to the Agreeableness cluster; "*Feel others emotions,*" "*Don't understand people who get emotional,*" "*Seldom get emotional,*" "*Am not easily affected by my emotions,*" "*Rarely notice my emotional reactions.*" The two items from O3 which were retained in the Openness cluster centred on cognition about emotional life; "*Enjoy examining myself and life*" and "*Try to understand myself.*" Two items regarding stimulating environments and activity ("*Do not like concerts,*" and "*Like to visit new places*") moved to the Extraversion cluster. Finally, a significant change from the FFM alignment of facets was seen for Extraversion, with 14 items shifting out of Extraversion and into

Conscientiousness. All 10 items from the facet "activity level" (E4: e.g. "*Like to take it easy*" (reversed) and "*Am always busy*") shifted to Conscientiousness. Four items from the Extraversion "assertiveness" facet (E3 e.g., "*Take charge*" and "*Take control of things*") were also absorbed into the Conscientiousness cluster.

In order to test if (and how) the Honesty–Humility and other clusters of the six-cluster solution map onto the five-factor solution in this sample, scores on each cluster, and on each factor were derived for each subject, and the correlations among these scores computed (this correlation matrix is shown in Table 3). As can be seen, the six cluster Honesty–Humility cluster scores correlated with scores on five-factor domain scores for C (.61) and A (.52), but also O (−.32), E (−.38) and (weakly) with N (0.1).

4.6. Discussion

Study 1 tested whether spectral clustering would extract an interpretable and well-fitting six-cluster solution from IPIP NEO item data when the scale parameter was small enough to enhance sensitivity to thinly sampled, but strong, item associations. The results supported this hypothesis, with five clusters extracted using larger scaling values, but six clusters fitting better for smaller values of the scaling parameter. It is worth reiterating that changing the scale parameter will not necessarily alter the results from spectral clustering (e.g. in Study 2 below, the scale parameter turns out not to affect the results). When results do differ at different scales, this can provide insights into the structure of the data. In this case, the results suggest that the IPIP NEO does sample (albeit weakly) behaviours comprising a sixth domain of personality.

In terms of psychological content, the sixth cluster closely followed that of the Honesty–Humility domain described in the HEXACO model—perhaps surprisingly so given that we worked from an item bank chosen to model the NEO-PI-R, and which may not adequately sample the facets of Honesty–Humility (Ashton & Lee, 2005). The sixth cluster was populated with Agreeableness and Conscientiousness items, supporting the observation by Lee and Ashton (2004) that their Honesty–Humility factor loads primarily on NEO facets falling within the Agreeableness and Conscientiousness domains when only five factors are extracted. Ashton and Lee (2005) found the NEO-PI-R facets of Straightforwardness and Modesty were those most closely related to Honesty–Humility. These map onto the IPIP facets of Morality and Modesty. Thus the prediction that IPIP Morality (A2) and Modesty (A5) items would constitute the Honesty–Humility cluster was supported. In addition we found that the Honesty–Humility cluster loaded strongly on Cautiousness (8 items) and also included four Dutifulness items. The present data, then, accord with the idea that facets usually incorporated into Conscientiousness are part of Honesty–Humility. The results thus inform HEXACO theory to some degree. Finally, Ashton and Lee (2005) reported a small negative correlation (−.12) of Honesty–Humility with FFM Neuroticism. This also was supported, with Honesty–Humility being associated with gluttony (overindulgence) and splurge-spending items linked to N in five-factor models.

The six-cluster solution versions of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness remained similar FFM representation, with some notable and informative changes. In particular, most of the Openness to Emotion facet (O3) was moved out of Openness in the six-cluster solution. O3 items tapping compassion versus callousness moved to Agreeableness, leaving Openness more focused on curiosity about – and understanding of -emotion. For Extraversion, activity and assertiveness moved to Conscientiousness. As activity is often viewed (along with sociability) as a core aspect of Extraversion (Eysenck, 1998) this is a significant change, leaving Extraversion much more focused on sociability, warmth (Friendliness, Gregariousness, Cheerfulness)

Table 3
Correlations among five-factor and six-cluster variables.

		Five-factor solution scores					Six-cluster solution scores					
		N	E	O	A	C	N	E	O	A	C	H–H
5	N	1										
	E	−0.22	1									
	O	−0.09	0.26	1								
	A	−0.06	0.06	0.12	1							
	C	−0.24	0.01	−0.06	0.20	1						
6	N	**1.00**	−0.24	−0.10	−0.05	−0.23	1					
	E	−0.23	**0.99**	0.24	0.07	−0.04	−0.25	1				
	O	−0.06	0.24	**0.99**	0.10	−0.11	−0.08	0.23	1			
	A	−0.15	0.26	0.29	**0.94**	0.22	−0.15	0.26	0.26	1		
	C	−0.36	0.29	0.17	0.10	**0.86**	−0.37	0.23	0.11	0.22	1	
	H	0.10	−0.38	−0.32	**0.52**	**0.61**	0.11	−0.38	−0.34	0.31	0.20	1

Note: This table shows the correlation between subjects' scores on each factor from the five-factor solution and each cluster from the six cluster solution. Correlations of 0.4 or above are in bold. (N = Neuroticism; E = Extraversion; O = Openness to experience; A = Agreeableness; C = Conscientiousness; H = Honesty–Humility).

and Excitement-Seeking (for instance items about enjoyment of travel and concerts moved from Openness to Extraversion).

In summary, spectral clustering at smaller values of the scale parameter in Study 1 revealed six clusters. The sixth cluster was substantive and clearly followed the characteristics of HEXACO Honesty–Humility, suggesting that the IPIP NEO does sample this domain. This suggests that, in a battery containing more items designed to assess Honesty–Humility, this sixth cluster should emerge at all values of sigma. If this interpretation is correct, and the five-cluster solutions found at higher values of the scale parameter in Study 1 reflect the paucity with which the Honesty–Humility domain is sampled in the IPIP NEO, then, in a dataset containing items which load strongly on the Honesty–Humility domain, these five-cluster solutions should never emerge. Instead, six-cluster solutions should be ubiquitously obtained independent of scale. We test this prediction in Study 2 using a large sample of responses to the HEXACO instrument of Ashton and Lee (2010).

5. Study 2

5.1. Introduction

Study 1 indicated that IPIP NEO items include Honesty–Humility, which was reliably identified at smaller values of the scale parameter that deemphasize weaker correlations. In Study 2, we wished to test the prediction that, in a dataset designed to strongly sample Honesty–Humility, (i.e., the HEXACO questionnaire) six clusters would be obtained at all values of the scale parameter sigma. Moreover, if the items for Honest-Humility are removed, then we support for six clusters should be lost, and instead, only five clusters should emerge. We tested these predictions in an independent sample of subjects who had completed the HEXACO, as reported by Ashton and Lee (2010).

5.2. Method

Participants. A total of 1,128 subjects, mainly college students, were combined from various samples (see Ashton & Lee, 2010). The final sample consisted of 693 females (mean age 21.0 years) and 429 males (mean age 21.6 years) and six subjects who did not provide their sex (mean age 28.0). All data were used.

Measures. Subjects completed the 100-item HEXACO-PI-R, which includes six domain-level constructs: Honesty–Humility (H), Emotionality (E), eXtraversion (X), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O).

5.3. Results

Study 2 used the same spectral clustering algorithms as in Study 1. As previously, we optimised against the value of *k* (the number of clusters) that minimised misclassification error. Fig. 7 shows the percentage of classification error across 1000 runs for each value of sigma and *k*. As before, cells marked with stars highlight the minimum misclassification error for each value of *k*. In the HEXACO data, the optimal number of clusters (*k*), i.e., minimum item misclassification was six for all values of sigma explored. Table 4 gives numerical values of the misclassification rate and the standard error across sigma and *k*.

In study 1, we hypothesized that the IPIP NEO data contained weak traces of Honesty–Humility, which were only adequately detected when this information was amplified by choosing appropriately low values of sigma, and that a measure such as the HEXACO questionnaire replete with items marking Honesty–Humility would discover this 6th cluster independent of sigma. As can be seen graphically in Fig. 7, and in line with this hypothesis, support for six clusters was found at all tested values of sigma (best fitting solutions are starred and all favour 6 clusters).

An implicit corollary prediction follows: if a subset of the HEXACO data is analysed, removing all items from the Honest–Humility scale, then evidence for the sixth cluster should also be removed, and five clusters should be optimal at all values of sigma. Alternative outcomes such as splitting off one of the other domains into two clusters would imply that the clustering algorithm is biased to extract additional domains when run at low values of sigma, independent of evidence for such additional clusters.

To test this prediction we re-ran the spectral clustering analyses on HEXACO item data, first removing all Honesty–Humility items. The results accorded with prediction: For all values of sigma tested (0.4 to 1), fit was best at *k* = 5. Thus, with the Honesty–Humility items removed, five clusters were returned at all scales (see Fig. 8), indicating that simply operating clustering with high values of scale does not arbitrarily over extract clusters, or promote facet-level constructs to domain level.

5.4. Discussion

Study 2 indicated that HEXACO data contain six clusters, and this solution fitted best at all usable values of sigma (values above 0.4). In addition, removing the Honesty–Humility items from the scale prior to analysis also removed support for the sixth-cluster. Jointly, these findings support the hypothesis that six domains of personality generate a well-fitting model, and are required at all values of sigma when the domains relating to humility and
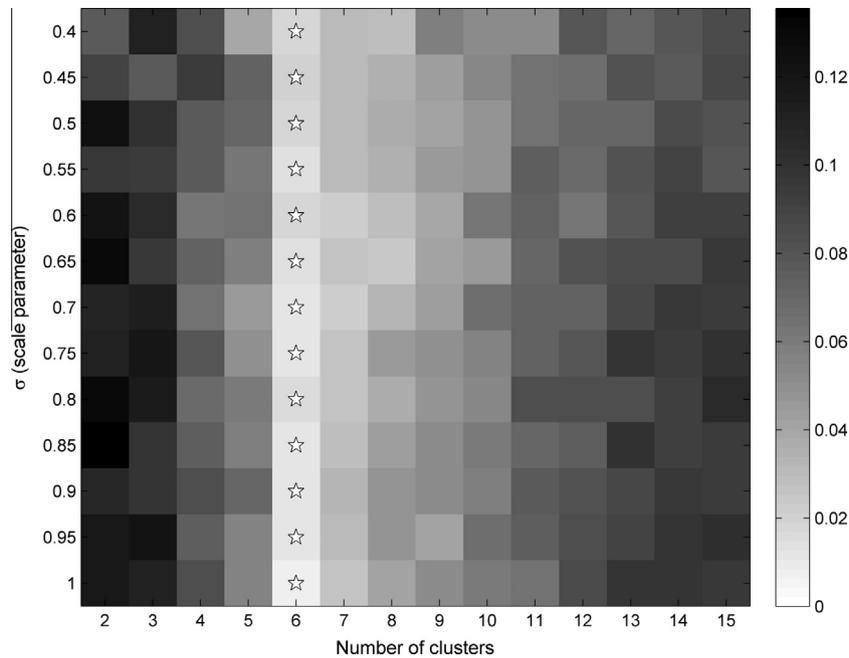
Fig. 7. Study 2: Mean of the minimum classification error percentages for all the trials for a given pair of k (the number of clusters) and σ (the scale parameter). As in Fig. 6, the stars indicate where the minimum occurs for each value of σ. In this study, minima are realised for six clusters over all σ.

Table 4
Classification error data from Study 2, showing mean misclassification errors (and SE) as a function of sigma and for each of 3–7 cluster solutions.

	# of clusters				
	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
Classification error when sigma = 0.4					
Misclassification rate	0.112	0.084	0.040	0.017	0.031
Standard error	0.011	0.009	0.006	0.003	0.004
Classification error when sigma = 0.5					
Misclassification rate	0.101	0.077	0.071	0.017	0.030
Standard error	0.010	0.009	0.006	0.003	0.003
Classification error when sigma = 0.75					
Misclassification rate	0.120	0.080	0.051	0.013	0.027
Standard error	0.012	0.008	0.006	0.002	0.003

Note: This table shows the mean and standard error of the proportion of items which were misclassified for different values of sigma and k. Each entry is based on 100 simulations. The standard errors are small enough to ensure that all the differences between the minimum rate and the other values within each pane are statistically significant at the 95% confidence level.

honesty are strongly sampled. The second set of analyses in Study 2 – removing the items of the Honesty–Humility domain – and analyzing the remaining subset of HEXACO data confirmed the prediction that, when Honesty–Humility is no longer sampled in a questionnaire where the items indexing the remaining five domains have been selected to avoid this domain, then six cluster solutions will notfit well, confirming that the analyses are not simply splitting factors at small values of sigma. We next discuss the results of both studies, and their implications.

6. General discussion

Four main results emerged. First, spectral clustering, even on NEO IPIP data, yielded a six-cluster solution corresponding to that of the HEXACO model, not only in having a sixth cluster, but in the Honesty–Humility-like nature of items comprising the sixth cluster, and in the effects on the remaining five clusters representing Extraversion, Neuroticism/Emotionality, Openness, Agreeableness

and Conscientiousness. By contrast, factor analysis yielded a six-factor solution containing a small and hard to interpret sixth factor. Despite its very different approach when compared to the algorithm of factor analysis, the five-cluster solution from the spectral clustering algorithm was highly similar to the five-factor solution from factor analysis (with both reflecting the FFM). The finding that six-cluster solutions were preferred depended on the weighting given to the weaker connections among the data (more weight favouring a five-cluster FFM solution, less weight favouring either a six-cluster HEXACO solution or a five-cluster FFM solution) casts light on the distinction between these two competing models for the structure of personality, and suggests that the NEO questionnaire contains traces of an Honesty–Humility dimension which is not detected by conventional factor analysis because it is too thinly sampled.

Second, spectral clustering suggested modifications to the psychological content of the 5FM domains, particularly in the six-cluster solution. Third, Honesty–Humility emerged at all scales in HEXACO data. Fourth, after removing Honesty–Humility items from the HEXACO, a 6th cluster could not be extracted at any value of scale, suggesting the content of the HEXACO items assessing the core 5FM domains is restricted to these domains.

7. Five cluster solutions

The five-cluster solutions reflects the domain structure postulated by the classic FFM. This is significant given the fact that these two statistical techniques – factor analysis and spectral clustering – are based on different transformations of the raw data, and have different optimisation targets. The present results, then, represent a robustness check on the FFM, which the FFM passes.

8. Psychological content of the six domains of personality

As might be expected for competing models of the same data, as can be seen in Table 2, the FFM and six-cluster HEXACO solutions are similar in many important ways. Each of the FFM domains, however, also was modified to a greater or lesser extent. The more substantial
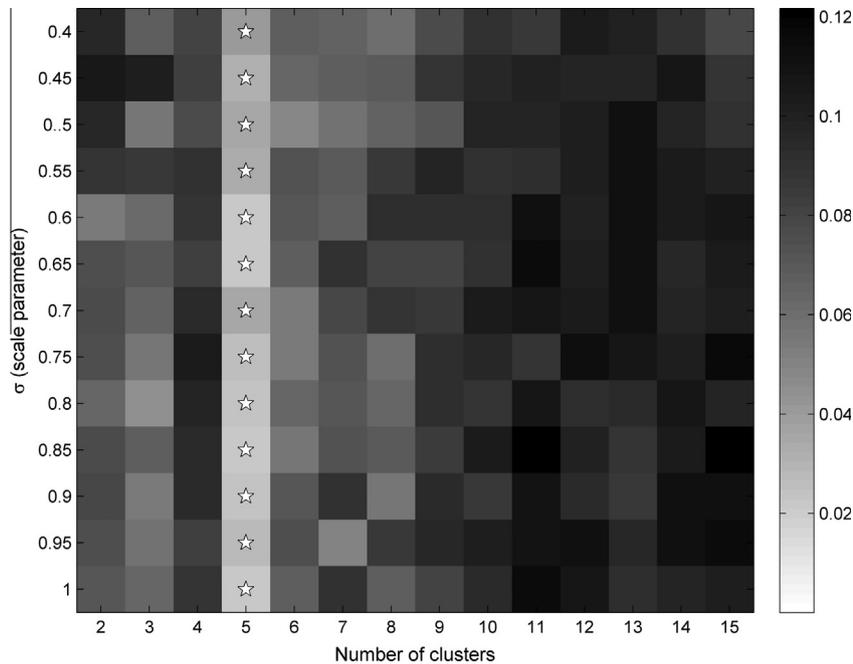
Fig. 8. Study 2 with Honesty–Humility questions removed: Mean of the minimum classification error percentages for all the trials for a given pair of k (the number of clusters) and σ (the scale parameter). As in Figs. 6 and 7, the stars indicate where the minimum occurs for each value of σ. In this study, minima are realised for five clusters over all σ.

of these are discussed next. Neuroticism retained most items losing content from N4 (self-consciousness) to Extraversion and from N5 (Immoderation) to Honesty–Humility and (for two items) to Conscientiousness. Openness retained most of its original items, but lost half the content of O3 (Openness to Emotions) in meaningful ways to Neuroticism/Emotionality. O3-items referring to the direct experience of emotion moved to Neuroticism, The affective content of Openness to experience, then, was refocused on understanding emotion, with O retaining more intellectual-items such as "*[I] try to understand myself*". Changes occurred also for Extraversion: this domain lost four items from E3 (Assertiveness) and the entire Activity Level facet (E4) to Conscientiousness. Extraversion was thus re-focused on warm sociability and excitement seeking, with Conscientiousness gaining an emphasis on capability for sustained activity, and social leadership. The implications of this change for Conscientiousness are discussed below.

Agreeableness and Conscientiousness were clearly the domains most likely to shift in focus when Honesty–Humility is extracted as a separate domain. Agreeableness changed in ways consistent with the theoretical predictions of the HEXACO model (Ashton & Lee, 2005). 6-cluster agreeableness lost all items relating to Modesty (A5) as well has half the items relating to Morality (A) and Cooperation (A4). These all shifted to Honesty–Humility. Together with its gain of empathy-related content from facet O3, Agreeableness was re-focused tightly on pro-sociality, cooperation, and fellow-feeling. The movement of hostility items from Neuroticism to Agreeableness predicted by Ashton and Lee (2007) did not occur. Hostility and Neuroticism, then, appear robustly linked, while Anger is more accurately reflects the affect associated with low Agreeableness.

As noted above, Conscientiousness gained the E3 (Assertiveness) and E4 (Activity Level) items from Extraversion, but lost around half of the C3 (Dutifulness) and most (8 of 10) items assessing Caution (C6). These moved to Honesty–Humility, and Conscientiousness was refocused away from caution and integrity and towards work capacity and task engagement, as predicted by Ashton and Lee (2007). This modified Conscientiousness

factor may more effectively absorb work-capacity related scales such as Need for Achievement (Duckworth, Peterson, Matthews, & Kelly, 2007; Tellegen, 2000). FFM Conscientiousness correlates highly with measures of effortful persistence (>.65) (Duckworth & Quinn, 2009), but it would be valuable to test whether scores on the conscientiousness cluster extracted in Study One are more strongly linked to persistence and achievement. Finally, the relationship of the sixth-cluster to the conventional five-*factor* domains (Shown in Table 3) shed light on how Honesty–Humility largely sits within the space delineated by the FFM. Honesty–Humility emerges not only as high Conscientiousness and Agreeableness, but also (more weakly) higher Neuroticism and lower Openness to experience and Extraversion.

9. Strengths and weaknesses

Both studies used large sample sizes and comprehensive item batteries. In addition, the batteries differed in ways that allowed us to test the robustness of the 6-cluster model in item pools chosen to represent five– and six-factor models. While this dataset is very large, and we were able to recover and validate solutions using resampling techniques within this parcel of subjects, it would be valuable to examine the results of spectral clustering in data derived from other cultures and subject pools, test formats, and, particularly, on the NEO-PI-R (Costa & McCrae, 1992) itself. External validation in terms of differential validity of the five- and six-cluster solutions in genetic and experimental studies would also be valuable.

10. Conclusions

Applying spectral clustering to two large personality datasets yielded six-cluster solutions in both IPIP NEO and HEXACO data. Scale effects indicated that the IPIP NEO questionnaire is comprehensive, in that it samples Honesty–Humility, but weakly, so that this domain is not detected by factor analyses, but is reliably

recovered by spectral clustering. The analyses also shed light on the psychological content of the personality domains, and suggest new associations based on these reformulations of psychological focus.

Acknowledgments

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jrp.2015.05.003.

References

Ashton, M. C., & Lee, K. (2005). Honesty–humility, the big five, and the five-factor model. *Journal of Personality, 73*(5), 1321–1353. http://dx.doi.org/10.1111/j.1467-6494.2005.00351.x.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review, 11*(2), 150–166. http://dx.doi.org/10.1177/1088868306294907.

Ashton, M. C., & Lee, K. (2010). Trait and source factors in HEXACO-PI-R self- and observer reports. *European Journal of Personality, 24*(3), 278–289. http://dx.doi.org/10.1002/Per.759.

Braun, R., Leibon, G., Pauls, S., & Rockmore, D. (2011). Partition decoupling for multi-gene analysis of gene expression profiling data. *BMC Bioinformatics, 12*, 497. http://dx.doi.org/10.1186/1471-2105-12-497.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245–276.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the big five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology, 66*(1), 93–114.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO personality inventory-professional manual*. Florida: Psychological Assessment Resources Inc.

Costa, P. T., Jr., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: Examples from conscientiousness. *European Journal of Personality, 12*(2), 117–134.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. http://dx.doi.org/10.1037/0022-3514.92.6.1087.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (Grit-S). *Journal of Personality Assessment, 91*(2), 166–174.

Eysenck, H. J. (1998). *Dimensions of personality*. New Brunswick, NJ, USA: Transaction Publishers.

Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the Psychoticism scale. *Personality and Individual Differences, 6*(1), 21–29.

Gignac, G. E., Bates, T. C., & Jang, K. L. (2007). Implications relevant to CFA model misfit, reliability, and the five-factor model as measured by the NEO-FFI. *Personality and Individual Differences, 43*(5), 1051–1062. http://dx.doi.org/10.1016/j.paid.2007.02.024.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216–1229.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe, 7*, 7–28.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*(1), 84–96. http://dx.doi.org/10.1016/J.Jrp.2005.08.007.

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14*(3), 332–346. http://dx.doi.org/10.1177/1088868310361240.

Horn, J. L. (1965). A rationale and a test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185.

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from Web-based personality inventories. *Journal of Research in Personality, 39*(1), 103–129. http://dx.doi.org/10.1016/J.Jrp.2004.09.009.

Joreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*(2P1), 183–202. http://dx.doi.org/10.1007/Bf02289343.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research, 39*(2), 329–358. http://dx.doi.org/10.1207/S15327906mbr3902_8.

Leibon, G., Pauls, S., Rockmore, D., & Savell, R. (2008). Topological structures in the equities market network. *Proceedings of the National Academy of Sciences of the United States of America, 105*(52), 20589–20594. http://dx.doi.org/10.1073/Pnas.0802806106.

Marsh, H. W., Ludtke, O., Muthen, B., Asparouhov, T., Morin, A. J., Trautwein, U., et al. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*(3), 471–491. http://dx.doi.org/10.1037/a0019227.

McCrae, R. R., & Costa, P. T. Jr., (1997). Personality trait structure as a human universal. *American Psychologist, 52*(5), 509–516.

McCrae, R. R., & Costa, P. T. Jr., (2003). *Personality in adulthood* (2nd ed.). New York, NY, US: Guilford Press.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175–215.

Ng, A. Y., Jordan, M. I., & Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. *Advances in Neural Information Processing Systems, 14*, 849–856.

Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research, 14*, 403–414.

Sammon, J. W. (1969). A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers, 18*, 401–409.

Spearman, C. (1927). *The abilities of man*. New York, NY: Macmillan.

Tellegen, A. (2000). *Manual for the multidimensional personality questionnaire*. Minneapolis: University of Minnesota Press.

Tiliopoulos, N., Pallier, G., & Coxon, A. P. M. (2010). A circle of traits: A perceptual mapping of the NEO-PI-R. *Personality and Individual Differences, 48*(1), 34–39. http://dx.doi.org/10.1016/J.Paid.2009.08.010.

Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing, 17*(4), 395–416. http://dx.doi.org/10.1007/s11222-007-9033-z.