

Convergence of Internal and External Structure for the California Child Q-set

David Bimler¹, John Kirkland¹, Hiram E. Fitzgerald²,
and Robert A. Zucker³

¹Massey University (New Zealand)

²Michigan State University (USA)

³University of Michigan (USA)

The language of personality traits includes single-word trait descriptors, and longer phrases or sentences. Evidence has accumulated that abstract, semantic relationships among single words have the same underlying structure as the empirical relationships when words are applied to individuals. The present study examines whether these two kinds of structure are also isomorphic for longer trait descriptors. Empirical descriptions and judgements of semantic similarity were collected among the descriptors comprising the California Child Q-set, or CCQ, and analysed with multidimensional scaling. Canonical correlation showed the solutions to be closely related to one another, and to independent sets of ratings available for the CCQ items. Informants' similarity judgements were not affected by the context in which they were made. The dominant dimensions of the solutions reproduce dimensions found previously for the single-word personality lexicon, indicating the two trait-descriptive languages to be closely parallel.

Keywords: big five, trait perception, internal structure, multidimensional scaling, California Child Q-set.

El lenguaje de rasgos de personalidad incluye a descriptores de rasgo de una sola palabra y frases u oraciones más largos. Hay evidencia acumulada de que las relaciones semánticas abstractas entre estas palabras tienen la misma estructura subyacente que las relaciones empíricas cuando las palabras aisladas se aplican a individuos. Este estudio explora si estas dos clases de estructuras son también isomorfas para descriptores de rasgo más largos. Se registraron las descripciones y los juicios empíricos de similitud semántica de los descriptores incluidos en el California Child Q-set (CCQ) y se analizaron con escalamiento multidimensional. La correlación canónica mostró que ambas soluciones están estrechamente relacionadas entre sí y con conjuntos independientes de puntuaciones disponibles para los ítems del CCQ. Los juicios de similitud de los informantes no estuvieron afectados por el contexto en el que fueron hechos. Las dimensiones dominantes de las soluciones reproducen otras dimensiones encontradas previamente en el léxico de personalidad de palabras únicas, indicando que los dos lenguajes de descripción de rasgos son cercanamente paralelos.

Palabras clave: cinco grandes, percepción de rasgo, estructura interna, escala multidimensional, California Child Q-set

This research would have been impossible without the informants, including students who found the task of arranging words into groups more congenial than formal lessons; the school principals who gave them the choice; or the assistants with the work of data collection, particularly Sarah Foley of Oakland University (MI). Reading the handwriting of informants in many cases called for advanced cryptanalytic skills, so special thanks are owed to Robyn Surville and Lynne McMorran who wrestled with loosely-structured data-collection forms as they entered GOPA-sorting data. Brian Vaughn and Alex von Eye were generous with advice and discussion. Jack Block provided access to expert criterion sorts. The paper has benefited considerably from the responses from two anonymous reviewers of an earlier version. Michigan Longitudinal Study data were collected with the support of National Institute on Alcohol Abuse and Alcoholism grant R37 AA07065 to Hiram E. Fitzgerald and Robert A. Zucker.

Correspondence concerning this article should be addressed to David Bimler, School of Arts, Development and Health Education, MUCE, Massey University, Private Bag 11-222, Palmerston North, (New Zealand). E-mail: d.bimler@massey.ac.nz

John Kirkland is currently at Royal University for Women, Bahrain.

Such is the human talent for linguistic invention that if some aspects of personality variation is important in human interaction and often talked about, then single-word descriptors should have entered the language to allow us to signal its presence or absence. In essence, this is the “lexical hypothesis” (Cattell, 1943). It appeals greatly to personality psychologists, since it implies that the language we speak is a distillation of psychological observations accrued across generations, so that a framework for personality description can be gleaned from assiduous study of the dictionary and Roget’s *Thesaurus*. In consequence, a series of publications have confronted the vocabulary of personality, seeking to reduce it to a representative but manageably-sized lexicon, suitable for specifying an individual’s position along the dimensions of ‘trait space’ (e.g. Goldberg, 1992; Peabody & Goldberg, 1989).

Other researchers have argued that some important facets of personality can only be specified by entire phrases (e.g. Block, 1961). If one-word descriptors exist for such aspects, they are too obscure or technical to enjoy wide currency. Conversely, the lexical hypothesis predicts that single-word descriptors and longer phrases occupy the same space and vary along the same dimensions, so that the choice of which to use for diagnostic or research purposes becomes one of convenience (Briggs, 1992), because both capture the same information. One difference is that it is difficult to determine how well an inventory of polylexemic trait descriptions samples the personality domain, since the pool of possible descriptors is open-ended, in contrast to the single-word vocabulary.

A secondary goal of the present study is to test the lexical hypothesis by comparing a semantic ‘map’ of single-word descriptors (Bimler & Kirkland, 2007) with a similar map representing the 100 sentence-length descriptors that comprise the California Child Q-set or CCQ (Block & Block, 1980). The comparison will concentrate on the spatial dimensions of this empirical map, but the distribution of items within it is also of interest.

The primary question addressed here is whether ‘internal’ and ‘external’ structure are the same for a list of polylexemic personality descriptors (the CCQ). ‘Internal’ structure underlies subjective judgements of semantic similarity among descriptors, while ‘external’ structure emerges from the inter-item correlations derived from their application in practice (Wiggins, 1973). In other words, we ask whether similarities and correlations are manifestations of the same trait space.

In the case of single trait words, several studies have found convergence between internal and external structures (e.g. Bimler & Kirkland, 2007; Peabody & Goldberg, 1989; Sneed, McCrae & Funder, 1998); that is, between the structures of similarity of meaning, and practical usage. This convergence is a corollary of the

lexical hypothesis: if the personality lexicon is simply a generalisation drawn from collective observation, then the semantics of any two words should reflect the empirical co-occurrences of the corresponding traits. The converse is also true. Within its narrow compass, the personality lexicon is a language, with internal semantic constraints which we expect speakers to follow when they use the lexicon to describe themselves or their peers. These constraints take the form of implicit rules governing the validity of applying two terms to the same target (rules which lexicographers have subsequently expressed explicitly in the form of definitions); the correlation in the use of those terms is thus influenced by the constraints. We are not proposing that semantic considerations are a Procrustean bed, deforming our perceptions of personality and behavior into their own mould – only that they impose a structure on how those perceptions are described.

The arguments for expecting an external/internal convergence are not so compelling when one turns to polylexemic trait descriptors. Empirical evidence one way or the other is scanty; studies of their external structure (noted above) are not matched by corresponding scrutiny of their internal structure. We are only aware of Sneed, McCrae and Funder (1998), whose Study 2 involved 64 brief phrases summarising behavior. To probe their internal structure, informants were provided with descriptions of the “Big-five” factors and asked to rate items on ‘diagnosticity’ for each one – in effect, they were arranging items along semantic gradients to indicate their semantic proximity to factor poles. These lay judgements showed an overall trend of agreement with external structure in a trait-attribution task.

In the two studies reported here, subjects provided estimates of inter-item similarity and covariance by following a variety of sorting procedures. These estimates were analyzed with multidimensional scaling or MDS, converting them into a geometrical representation of the items as points in a many-dimensional ‘trait space’. Points are positioned so that distances between them reflect the dissimilarities between the corresponding items. The advantages of non-linear MDS include parsimony and clarity: it may fit a two-dimensional solution to data which require six factors in a FA solution (Schlesinger & Guttman, 1969). MDS has been applied to the 57 items of Form A of the Eysenck Personality Inventory (Green & Walkey, 1980), and to a battery of bipolar trait-adjective scales designed to quantify the FFM (Maraun, 1997). Both studies found that a two-dimensional ‘map’ accounted for the pattern of inter-item correlations.

From the present data, we obtained five matrices of internal-structure similarity for the CCQ, for comparison with two matrices of external structure estimates – inter-item correlations derived from empirical application of the items. Separate MDS solutions were compared. If

they prove to be compatible, the various matrices can be combined into a single consensus map of trait space, and we will be able to pose the secondary research question noted earlier: how far that map's dimensions correspond to those found in MDS studies of single-word descriptors (e.g. Bimler & Kirkland, 2007).

It may be that the MDS representation as points in a trait space is a better approximation for longer descriptors; single words might be more akin to clouds, sometimes straddling the boundaries between factors, sometimes impinging on one another. Phrases and sentences make up in precision what they lack in concision; they are circumscribed and qualified whereas single words are burdened with multiple, flexible, overlapping meanings (polysemy), exacerbated by the lack of consensus about their definitions.

The CCQ is a version for children of the California Q-set or CQS. This notable selection of polylexemic personality predicates or descriptors consists of 100 statements about traits and behavior, intended as a descriptive language in which an adult's personality can be discussed by indicating how accurately or inaccurately each statement applies to him or her (Block, 1961). It aims to be a comprehensive concourse, covering all major facets of personality without restricting itself to the constructs of any single model of psychological function. Within this constraint, the CQS gives considerable coverage to the subtleties of ego-control and ego-resilience, thought by its compilers to be central to psychological dynamics. Another instance is the NEO Personality Inventory. This was designed to operationalise the 'Five-Factor Model' of personality, the FFM (originally derived from single-word factor-analytic studies, in accordance with the lexical hypothesis). Factor analysis (FA) has been applied to self-report and peer-description data from the CQS (Lanning, 1994, McCrae, Costa & Busch, 1986); the California Adolescent Q-set (Lorr, 1978); the Myers-Briggs Type Indicator (McCrae & Costa, 1989); and the NEO-PI (McCrae & Costa, 1987). In many cases the resulting factors can be identified with those underlying the FFM.

We chose the CCQ because several independent forms of information are available as ways of testing the validity of the resulting trait space, and as guides for interpreting its dimensions. These forms include tables of factor loadings from factor analyses; correlations between scores for children on some objective index and the values assigned to particular CCQ items; and criterion sorts. A criterion sort is a characterisation of a specified abstraction or theoretical construct such as Social Desirability (Waters, Noyes, Vaughn & Ricks, 1985). In each sort, a panel of informants, familiar with the construct in question, have gone through the items and quantified each one's applicability to or degree of association with the construct, via a 'Q-sorting' process described below.

Criterion sorts (sometimes known as prototypes or templates) can be regarded as semantic judgements, i.e. as another manifestation of internal structure. Conversely, another form of external structure is an 'empirical template' in which the numerical value of each item is its correlation with some index of an operationalised construct. In further tests of the generality of the combined MDS solution, we compared it with both forms of information, and also with other authors' factor-analytic conclusions about the external structure of the CCQ.

Completeness of the CCQ

Several existing personality instruments were designed to operationalise specific models. Often the compilation of a trait inventory is only the first step, followed by a process of enhancing its deduced or theoretical structure of latent variables (factors), and maximising the consistency of the measurement scale associated with each factor. Intermediate or interstitial items which load significantly on more than one factor are winnowed to make room for 'marker' items which load highly on a single factor (Goldberg, 1992). Multidimensional scaling of the adjusted item set reveals a tightly-clustered arrangement (e.g. Bimler & Kirkland, 2007) in which factor poles are isolated by these artificial or exaggerated voids. In contrast, the CCQ was designed to be broadly inclusive, so that theory-driven expectations would not restrict its capacity to capture data. Although FA reveals factors comparable to those of the FFM (van Lieshout & Haselager, 1994), the CCQ should also contain 'bridges' of items intermediate to the factor poles, making it well-suited to examining any connections between these nominally independent factors.

Guttman (1966) coined the term 'circumplex' to label a distribution of points, representing items in a spatial model, which is circular as well as continuous, so that a point's angular coordinate suffices to describe that item. Here we also apply the term to higher-dimensional analogues, such as a spherical shell in three dimensions. It would come as no surprise if a circumplex provided a good representation of the similarities and correlations among the items of the CCQ. One school of thought looks for circumplex structure in the domain of personality traits, or at least the sub-domain of interpersonal behavior (e.g. Gurtman & Pincus, 2000). Hofstee, de Raad and Goldberg (1992) fitted the trait lexicon into a five-dimensional circumplex model.

However, one must bear in mind the possibility that circumplex structure in a MDS solution has been imposed or exaggerated by the methodology. Noisy or low-resolution data can produce an 'annulus artefact', displacing points in a solution towards a circular or spherical annulus (Goodhill, Simmen & Willshaw, 1995). The converse is also possible – that MDS analysis might

exaggerate some slight clustering or non-homogeneity among the items.

To address these concerns we collected additional data in Study 2, this time with nearly half the items missing, breaking the inclusiveness of the CCQ. The omissions were not random or evenly distributed within the CCQ territory; particular item clusters were retained. The question is whether these omissions will be apparent, as gaps or voids within the MDS solutions for the partial item set. If so, we can infer that the context of other items has little effect on judgements of inter-item similarity (response style); and that any tendency towards a circumplex observed for the complete item set is not necessarily an artefact of MDS methodology. We will be in a position to tell how well the CCQ met its target of inclusiveness. Any void in the complete-set MDS solution will presumably correspond to an actual gap in the coverage of the CCQ, reducing its capacity to measure that facet of personality.

Study 1: Complete Set

Method

Stimuli. The original California Child Q-set (Block & Block, 1980) is an age-appropriate form of the widely-used California Adult Q-set. It consists of 100 statements about personality, cognitive and social characteristics of children. This study involves data collected for the original version, here labelled as the CCQ(o), and also for a revised version, the CCQ(r). Caspi et al. (1992) created a 'Common-language' CCQ by rewording 89 items to increase their readability. They expanded technical terms into language that lay sorters could understand, extensively consulting personality psychologists to ensure the revisions did not distort the meanings of the items. The CCQ(r) includes further revision, with minor changes made to eliminate male-only language while bringing the wording closer to the New Zealand vernacular.

Items were printed on slips of thin card measuring 75 by 35 mm.

Participants. Three groups of university-age students were recruited. Thirteen students from a New Zealand College of Education (without any specialized knowledge of psychology) took part as *Sample 1*, receiving a token NZ\$5 payment for the hour-long session. This sample also included 64 secondary-school students, aged from 13 to 15, and roughly balanced by gender.

Thirty-one psychology students from Oakland University (MI.) took part as *Sample 2*, receiving course credit. Four human-development graduate students provided semantic gradients as *Sample 3*.

The observers providing Q-sorts as *Sample 4* were trained clinicians (graduate students in clinical or educational psychology), ranging in age from 23 to 33.

Procedures. To make a proximity judgement requires the simultaneous consideration of at least two items, while it is often convenient to employ procedures that present all items simultaneously, so that informants' decisions are made in the context of the entire item set. Sorting items into groups is one example of this (e.g. Church & Katigbak, 1989; Sneed et al., 1998). Our research used variants of the similarity-sorting method. The variations in procedure reflect improvements made in the course of our research, to collect more data from each informant.

Participants in Sample 1 followed a three-stage procedure that we have called GPA-sorting, for Grouping, Partition, Addition. The Grouping stage consisted of arranging items into groups on the basis of perceived similarity, using participants' own criteria as to what constitutes 'similarity'. The number of items per group was unrestricted (single-item groups were permitted); a range of 10 to 20 was suggested as a suitable number of groups. Next, to provide an insight into the arrangement of items within the groups, participants were instructed to create and record a finer subdivision (i.e. a partition with more groups). They did this by inspecting each of the groups they had initially created, deciding whether the items comprising it were homogeneous in meaning, and if not, how it could be broken into subgroups. After this Partition stage, they restored the original groups and proceeded to the Addition stage. This consisted of merging the groups in a series of steps (selecting the two most similar groups at each step and combining them into one), until further integration was not possible (Sherman, 1972). The membership of each group was recorded after the initial grouping and after each subsequent merging step.

The information from GPA-sorting primarily describes the small proximities between similar items. A GOPA-sorting task complemented this with information about large dissimilarities, obtained in an 'Opposite-sorting' stage, after the G-sorting stage. Sorters were instructed to choose the two groups out of those they had constructed that provided the strongest contrast or antinomy, i.e. the pair which were most opposite in meaning. Putting those groups aside, they could choose a second pair of opposites, up to three times. They then restored the groups in preparation for Partition-sorting. A secondary purpose of this procedure was to discourage the merging of diametrically-opposite item-groups in the A-stage. Otherwise, because such exact antonyms share a common theme, they can be seen as similar.

Target descriptions were elicited with Q-sorting, the standard procedure for the CCQ. As with the similarity sorting, this procedure sets the items in context by presenting them simultaneously while they receive numerical values (in contrast to the usual rating techniques where each word or phrase or sentence is typically rated independently of other items in the inventory). Participants are asked to sort

the items into piles according to how well they describe a target. The number of piles is set by the researcher; they form a sequence ranging from 'most applicable' items to 'least applicable'. The number of items in each pile is usually predetermined (Block, 1961).

The value assigned to the i -th item is the number of the pile in which it was placed, ranging from -4 ('most applicable') to +4 ('least applicable') in Sample 4, and from -3 to +3 in Sample 5 (see Study 2). It is convenient to label these values as v_{ki} , where k is an index identifying the target description. To label the k -th Q-sort on its own (written as a row vector) we write v_k .

Sample 1. Thirteen university-age students applied the GPA procedure to the CCQ(r) items. They shuffled the deck into random half-decks to be sorted separately, yielding 26 GPA sequences. Sixty-four school students followed the same procedure, also with the CCQ(r) set. They created smaller subdecks with about 30 items in each. Some students sorted two of these third-decks in the time available, while others only managed one, eliciting 97 sorting sequences in total.

Sample 2. Thirty-one university students applied the GOPA procedure to the CCQ(o) items. They shuffled the items into random half-decks for sorting, yielding 62 GOPA sequences.

Sample 3. A preliminary hierarchical-clustering analysis of the data from Sample 1 arranged the CCQ(r) items in a tree structure (dendrogram). Forty-two items were selected, distributed evenly through the dendrogram. Four graduate students in human development rated these items on a scale of 1 to 5, on their degree of association with five abstractions (Ego-resilience, Ego-control, Activity, Social Desirability, Locus of Control).

Sample 4. 529 Q-sorts were collected in the course of the Michigan Longitudinal Study (Fitzgerald, Zucker & Yang, 1995; Zucker et al., 2000), using the CCQ(o) to describe 202 boys and 42 girls, aged 3-5, and 212 boys and 73 girls, aged 9-11. Each Q-sort was completed immediately after a full psycho-social assessment of the child which included parent and teacher ratings, as well as direct observation of the child in the child's home. A nine-pile template was used with the number of items per pile following a quasi-normal distribution (5, 8, 12, 16, 18, 16, 12, 8, 5).

Analysis

We converted the raw data into four 100-by-100 matrices of estimated inter-item proximity as follows:

- D[Sr], from the G, P, A stages of Sample 1, for CCQ(r) items;
- D[So], from the G, P, A stages of Sample 2, for CCQ(o) items. Individual entries were calculated as described in Bimler and Kirkland (2007, Appendix A).

- D[O], from the O-stage data of Sample 2. Individual entries were calculated according to the algorithm in Bimler and Kirkland (2007, Appendix B).
- D[Q], from the targeted Q-sorting of Sample 4. Each entry is the product-moment correlation between values of the i -th and j -th CCQ(o) items across 529 Q-sorts.

The goal of MDS is to arrange points in a P -dimensional space so that distances between them reflect the inter-item proximities. That is, it represents items within a multidimensional 'space' or map. Note, however, that the optimum P is not immediately obvious. We applied non-metric MDS to each proximity matrix separately (Kruskal's algorithm), and examined values of P from 2 to 6. The coordinates of points in the resulting solutions can be written as 100-by- P matrices, labelled Sr_p , So_p , O_p , and Q_p .

The first three matrices were also analysed in conjunction, using the repeated-measures mode of MDS. Four- and five-dimensional 'group solutions' G_4 and G_5 were obtained (each one is a compromise or consensus across the three forms of similarity). Here G_4 is a 100-by-4 matrix; its i -th row specifies the spatial location of the i -th item, and contains the coordinates $(g_{i1}, g_{i2}, g_{i3}, g_{i4})$.

We checked the similarity of particular pairs of solutions by calculating two indices: (1) the product-moment correlation r between their respective sets of inter-point distances, and (2) the least-squares Procrustes statistic g_i between point coordinates, calculated with the program GPA. g_i indicates the irreducible discrepancy between configurations, as a fraction of total variance, when they have been rescaled and rotated to maximize the fit between them.

To determine the number of stable dimensions, we worked with the five-dimensional solutions to ensure that no structure of importance was lost, and applied Canonical correlation (CANCORR) to examine their mutual compatibility. CANCORR extracts a pair of linear combinations from the two coordinate sets under comparison, such that the correlation between them R_c is maximal. It can extract further pairs of linear combinations of coordinates, providing correlations R_2, R_3, R_4, R_5 - each new combination being independent of those previously extracted - where the significance of each successive correlation is tested using Wilks' Lambda statistic. The number of significant correlations indicates the number of mutually-recognizable dimensions shared between the two coordinate sets.

Each of the solutions Sr_5, So_5, O_5 was compared against the group solution G_5 , and then against Q_5 . Three other comparisons were performed - Sr_5 with So_5 , Sr_5 with O_5 , So_5 with O_5 - to confirm that the data matrices had enough in common to justify combining them.

The ratings for 42 items elicited from Sample 3 were averaged across the four informants, resulting in five

'criterion sorts'. We write w_{fi} as shorthand for the mean value for the i -th item on scale f ($1 \leq f \leq 5$). These values provide a further validity test of G_4 in which they serve as the dependent variables in a series of multivariate regression calculations. The purpose of each analysis is to account for w_{fi} – for each f in turn – in terms of an optimal combination of four independent variables, the item coordinates. The multivariate correlation R indicates the compatibility of that criterion sort with the internal structure of the CCQ. These values will also help us to interpret the dimensions of the four-dimensional solution G_4 , by correlating the w_{fi} with the items' coordinates. In the context of the geometrical framework, multivariate regression can be considered as a search for a four-dimensional direction or vector in G_4 , running through the solution from the side where items with low values of the dependent variable w_{fi} are located, through to the side with the highest values.

The roles of validity test and interpretation are not confined to the ratings from Sample 3. The following information is also available from the literature, and amenable to the same analyses involving multivariate regression and pairwise correlations. Like the current ratings, data sets (a) and (b) are internal data. The other sets are external data.

- a. Criterion sorts characterising the abstractions Ego Resilience, Ego Undercontrol and Field Independence (Block & Block, 1980).
- b. Criterion sorts for Social Competence, Self Esteem and Social Desirability (Waters *et al.*, 1985, Table 1).
- c. Loadings on each of seven Varimax-rotated factors emerging from factor analysis of CCQ data (van Lieshout & Haselager, 1994, Table 15.2);
- d. Each item's empirical association with Ego Resilience (Block & Block, 1980; Tables 2.3B, 2.5). This index of association is the correlation between the value assigned to that specific item in CCQ descriptions of three- or four-year-old children, and the same child's composite score on experimental tests of Resilience.
- e. Each item's empirical association with Ego Undercontrol, for three- and four-year-olds (Block & Block, 1980; Tables 2.3A, 2.5).
- f. Each item's association with 'Visual Attention', i.e. the amount of time that a given child was attracting the gaze of other children (Vaughn & Martino, 1988, Table 2; Waters, Garber, Gornal & Vaughn, 1983, Table 3).
- g. Each item's association with 'Activity', i.e. the amount of time that a child was physically active, measured objectively by a wrist-worn actometer (Buss, Block & Block, 1980, Table 2). These data were acquired for three-, four- and seven-year-olds.
- h. Each item's empirical association with tests of Field Independence (Kogan & Block, 1991, tables 10.1 and 10.2).

Results

The CANCELL comparisons indicated that four dimensions were generally replicated across various forms of data (at $p < 0.01$). The solution derived from O-sorting data has only two dimensions that can be recognized in other solutions; others are presumably obscured by the higher noise level in O-sorting distance estimates. In particular, four dimensions were mutually recognisable between the external structure of Q_5 and the combined internal structure of G_5 . We were able to interpret the first four dimensions of G_5 , whereas an interpretation for the fifth dimension eluded us. For these reasons, we opted to concentrate on four-dimensional solutions such as G_4 for subsequent analyses. Retaining four dimensions from MDS will also facilitate a comparison with the internal structure of trait adjectives (Bimler & Kirkland, 2007).

The small pool of Sample 2 informants introduces statistical fluctuations between Sr_4 and So_4 . Even so, the differences were not substantial ($g_7 = 0.18$ and $r = 0.76$). This confirms that the exact wording of the items had little impact on their perceived structure, not even in combination with a switch from naive to psychologically-sophisticated informants, so that combining the data sets is justifiable.

No substantive differences between external and internal structure were found by comparing G_4 with Q_4 , using distance correlation ($r = 0.80$) and the Procrustes statistic ($g_7 = 0.08$).

Two views of G_4 are shown in Figure 1. John, Caspi, Robins, Moffitt and Stouthamer-Loeber (1994) considered the Common-language version from the perspective of the Big Five and selected 48 CCQ items for which they could see a strong *a priori* association with the positive or negative pole of one or other of the factors. They also applied FA and found two additional unipolar factors, each epitomized by five high-loading marker items. In total John *et al.* selected 55 marker items (three of them serving to epitomize more than one factor). Symbols in Figure 1 indicate these markers. If this 'map' is meaningful, we expect it to segregate factor markers into distinct zones, and the poles of each bipolar factor to be diametrically opposite or nearly so.

The result of each multivariate regression calculation predicts the values of some dependent variable w_{fi} from the locations of items in G_4 , more or less accurately. The generally large multivariate correlations R (listed in Table 1) indicate that in most cases the predictions were accurate. This implies that ranking judgements about the items (by descriptive aptness for some abstraction) were consistent with similarity sorting; that is, the same kind of knowledge about the items was involved in both kinds of judgement. Table 1 also lists the correlations between these various forms of information about the items, and their coordinates in G_4 .

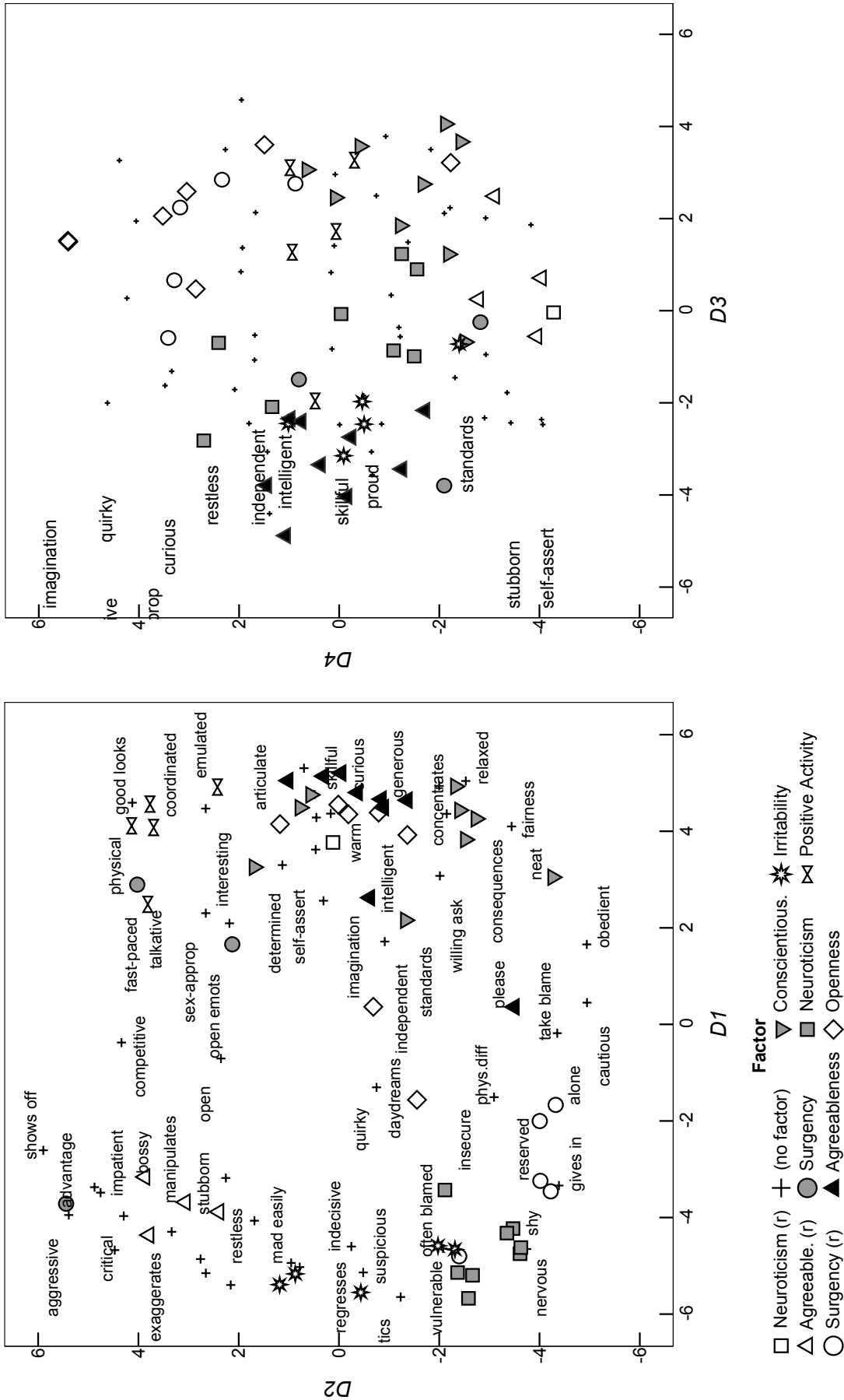


Figure 1

Figure 1. Locations of the 100 CCQ items in a four-dimensional MDS solution G4. (a) D1/D2 plane. (b) D3/D4 plane. Items that mark any of the “Big-5-plus-two” factors (John et al., 1994) are shown by corresponding symbols. Text labels for some of the points are very brief versions of the actual content of those items.

Table 1

Correlations between the coordinates of the CCQ items in G_4 and their values w_{fi} on various constructs or empirical profiles. N is the number of items for which values are reported. R is the multivariate correlation, when multivariate regression is used to account for the w_{fi} as combinations of the items' coordinates

	N	g_{i1}	g_{i2}	g_{i3}	g_{i4}	R	Source:
Ego-resiliency	100	0.91	—	—	—	0.917	Block & Block (1980).
Ego-undercontrol	100	—	0.61	—	0.20	0.678	
Field Independence	100	0.60	—	0.51	—	0.807	
Social competence	100	0.89	—	—	—	0.910	Waters <i>et al.</i> (1985).
Social desirability	100	0.90	—	—	—	0.907	
Self-esteem	100	0.86	—	—	—	0.887	
Ego resiliency	42	0.90	—	—	—	0.932	from Sample 3.
Ego control	42	—	-0.58	—	—	0.673	
Activity	42	0.45	0.52	—	—	0.708	
Value	42	0.88	—	—	—	0.919	
Locus of control	42	0.31	—	0.63	—	0.743	
F1, Agreeableness	100	0.74	-0.44	-0.30	—	0.907	van Lieshout and Haselager (1994).
F2, Emotional stability	100	0.66	0.44	—	—	0.800	
F3, Conscientiousness	100	0.60	—	0.42	—	0.793	
F4, Openness	100	0.68	0.20	—	0.25	0.766	
F5, Extraversion	100	—	0.64	-0.21	—	0.708	
F6, Motor activity	100	0.27	0.61	0.20	—	0.704	
F7, Dependency	100	—	—	-0.28	—	0.410	
Ego-Resiliency at 3	33	0.95	—	—	—	0.968	Block & Block (1980): correlates of experimental scores.
Ego-Resiliency at 4	67	0.90	—	—	—	0.915	
Ego-Resiliency	54	0.59	—	0.33	—	0.652	
Ego Undercontrol at 3	34	-0.42	0.91	—	—	0.918	
Ego Undercontrol at 4	63	-0.60	0.69	—	—	0.879	
Ego Undercontrol	54	-0.71	0.50	—	—	0.842	
Visual Attention 1	82	0.43	0.69	—	—	0.794	Waters <i>et al.</i> (1983).
Visual Attention 2	32	0.54	0.87	—	—	0.952	Vaughn & Martino (1988).
Actometer at 3	62	—	0.87	—	—	0.885	Buss, Block & Block (1980).
Actometer at 4	62	—	0.85	—	—	0.862	
Actometer at 7	62	—	0.82	—	—	0.828	
Field Independence	63	0.84	—	0.32	—	0.865	Kogan & Block (1991).

Correlations are omitted when $p > 0.05$; shown in bold when $p < 0.01$.

Crucially, empirical data – associations between CCQ items and particular objective measurements – could also be summarised as vectors or directions within the semantics-based model, indicating convergence of external and internal structures. The same analysis was applied to the factor loadings from van Lieshout and Haselager (1994), with their seven factors F1 to F7 serving in turn as the independent variable. Clearly F7 (“Dependency”) is not represented well in G_4 , but otherwise the convergence is good.

It will be instructive to clarify the external / internal convergence by interpreting the Big Five factors (external

structure) in terms of the dimensions of G_4 (internal). First, however, we must interpret these dimensions themselves. Recall that $\{g_{i1}, g_{i2}, g_{i3}, g_{i4}\}$ are shorthand labels for the coordinates of the i -th item.

Our analysis of the axes of trait space begins with the fourth dimension ($D4$). Otherwise this report will pay relatively little attention to $D4$, which is perhaps a distinction between traits from the intra-personal and interpersonal sub-domains of psychology. Another possibility is a polarity between Introversion and Extraversion – using these terms in the original Jungian sense, of internal focus versus external, social-world focus.

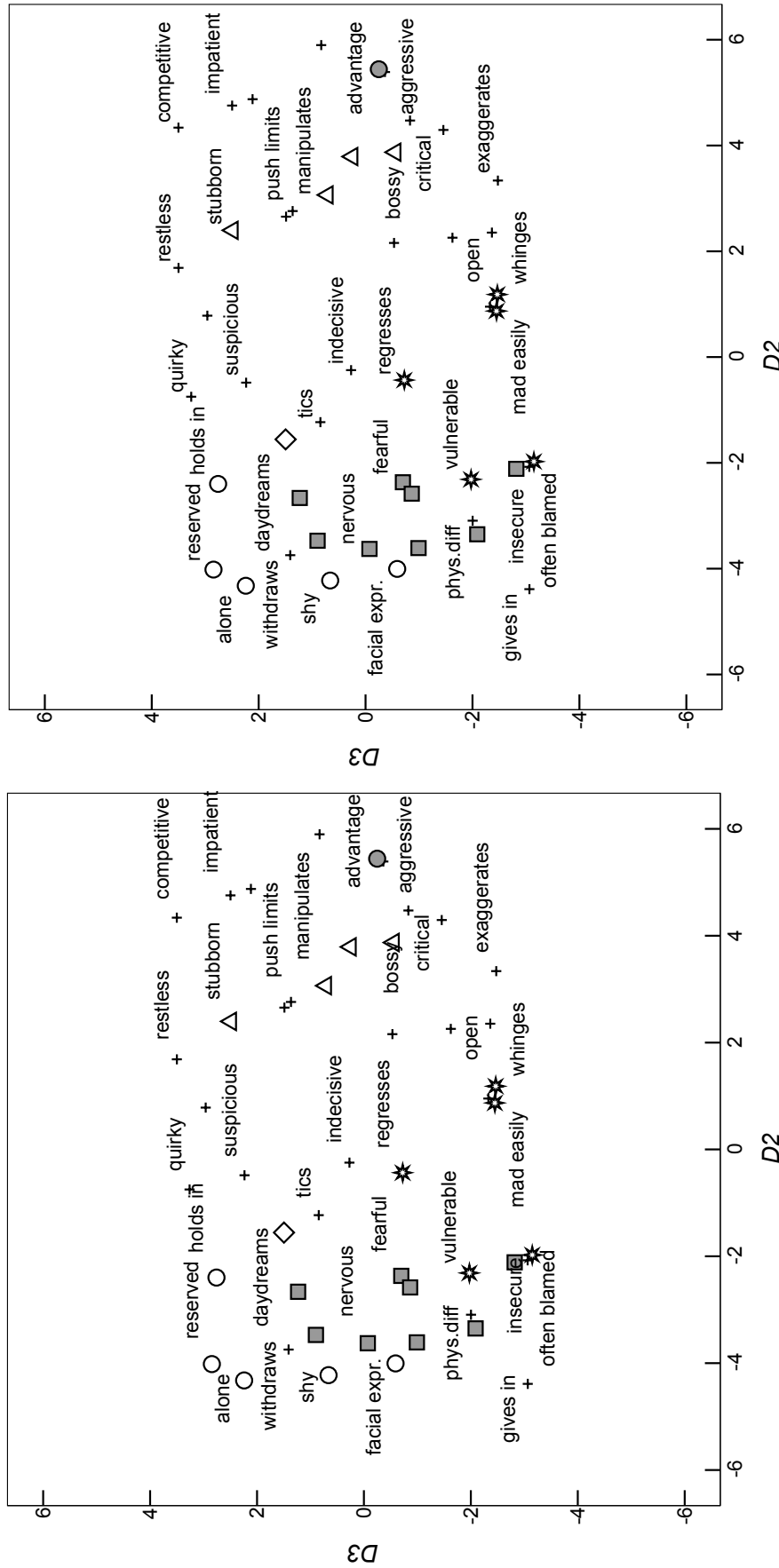


Figure 2. (a) D1+ hemisphere. (b) D1- hemisphere.

Figure 2. MDS solution G4 for the 100 CCQ items, projected onto the D2/D3 plane and labelled according to which of the “Big-5-plus-two” factors they mark (symbols as in Figure 1). (a) D1+ hemisphere. (b) D1- hemisphere.

This is consistent with the negative associations between *D4* and the empirical index of Visual Attention (which is clearly related to social interaction).

The first dimension (*D1*) separates socially-desirable and undesirable items. Note the high correlation between each item's position g_{i1} along *D1*, and its values of Social Desirability and Value (Table 1). At one extreme of *D1* were items such as "Tends to go to pieces under stress" and "Has rapid shifts in mood"; at the other were "Is cheerful" and "Develop genuine and close relationships" (Figure 1). It is possible to identify this axis with Evaluation, the first of three ubiquitous, fundamental factors which Osgood (1971) encountered in a range of semantic domains. Contributions from *D1* also dominate the clinical constructs of Ego Resilience, Self-Esteem and Social Competence, but none of these are evaluatively neutral: for instance, items with connotations of Ego Resilience are also desirable. There are few items in G_4 where g_{i1} is close to zero, and other dimensions can be seen in isolation. That is, most items have an evaluative content, one way or the other.

Apart from that discontinuity along the first dimension, the first three dimensions of G_4 turn out to arrange the items in a three-dimensional circumplex, i.e. a roughly spherical shell. That is, the points are roughly equidistant from the centre. Figure 2 plots the negative-Evaluation hemisphere ($g_{i1} < 0$) separately from the positive-Evaluation hemisphere ($g_{i1} > 0$), projecting both onto the *D2/D3* plane. In effect, each hemisphere is flattened onto the page. This introduces distortions for items near the join between the hemispheres (i.e. at $g_{i1} = 0$), but as already noted, there are few of these.

The high correlation between g_{i2} and objective correlates of 'activity' suggests that we identify the second dimension (*D2*) with Activity, a second fundamental factor in Osgood's schema. *D2* ranges from "Shy and reserved", "Tends to yield and give in" at one extreme to "Tries to be the centre of attention" and "Sees what s/he can get away with" at the other. Contributions from *D2* also dominate the constructs 'Activity' and (negatively) 'Ego control', as well as the empirical templates for 'Ego Undercontrol' and 'Visual Attention'. The items which epitomised the factor of Positive Activity (John *et al.*, 1994) all have high values of g_{i2} , combined with positive g_{i1} (i.e. connotations of favourable Evaluation).

The third dimension (*D3*) lends itself to such interpretations as "Individualism versus cooperation", "Independence" or "Self-regulation". At the positive extreme are items such as "Likes to be independent", "Does not give up easily". Accordingly, markers of Dependency can be found towards the negative extreme of the axis, such as 'Tries hard to please other people' and 'Gives in easily in conflict situations', along with items signalling amenability and social sensitivity. Table 1

shows a strong correlation between g_{i3} and two constructs related to independence and psychological autonomy, Field Independence and Locus of Control (though these constructs also include a substantial Evaluative contribution). The g_{i3} coordinates partly account for the items' associations with empirical indices of Field Independence, and their loadings on the Dependency factor identified by van Lieshout and Haselager (1994). In Osgood's three-factor schema, this dimension could be subsumed under the general rubric of Potency.

As noted above, John *et al.* (1994) selected 55 CCQ items to mark the Big-5 factors plus two additional unipolar factors. The locations of these markers in Figures 1 and 2 allow the factors to be described in terms of the three dimensions, as follows:

FI (*Surgency*) is aligned with Activity, with a generally positive Evaluative component.

FII (*Agreeableness*) is high on Evaluation, low on Potency (the FII- items marking the factor's negative pole are negative-Evaluation or Bad, but positive-Activity rather than high on Potency as symmetry would predict).

FIII (*Conscientiousness*) is Evaluative, and slightly Potent.

FIV (*Neuroticism*) is the converse of Emotional Stability. Its markers are Bad and Inactive: Inactive versions of FII-, as it were.

FV (*Openness to Experience*) is Potent, and generally high-Evaluation, though the markers extend over a range of *D1* values, and overlap with FIII markers.

FVI (*Irritability*) is Bad and negative on Potency.

FVII (*Positive Activity*) is Active and positive-Evaluation, with a connotation of positive Potency distinguishing it from Surgency.

There are few points in the map antipodal to the markers of FIII and FV, indicating an absence of specific markers for the negatives of conscientiousness and openness (noted also by van Lieshout and Haselager, 1994). The CCQ still allows these qualities in a child to be captured by assigning negative values to the items at the positive poles.

Finally, it is worth commenting on the two main differences between G_4 and the external-structure solution Q_4 (space does not allow an illustration of the latter, or a detailed discussion of its features). Firstly, the third- and fourth-dimension coordinates of the items are generally smaller in Q_4 than in G_4 . That is, items tend to congregate nearer to the *D1/D2* plane. One can infer that although positive or negative connotations of 'Independence' and 'internal focus' might contribute to an item's semantic content, conceptually distinguishing it from other items, these connotations are less important when the item is used in practice. Secondly, focussing on the *D1/D2* plane, Q_4 shows more continuity; it is a more complete

circumplex. The conceptual discontinuities in G_4 – the gap between desirable and undesirable high-Activity traits, and the similar gap between forms of low Activity – do not modulate the descriptive use of the items strongly enough to create corresponding gaps in Q_4 .

Study 2

Method

The stimuli were the 55 CCQ(r) items picked by John *et al.* (1994) as factor markers. We refer to this ‘minimal’ or ‘marker’ subset as the CCQ(m). Two samples of secondary-school students were recruited, aged between 13 and 15, and roughly balanced by gender.

Sample 5. Forty-two school students applied GPA-sorting to the CCQ(m) items, shuffling them into random half-decks for sorting. Only 80 GPA sequences were collected, because four participants ran out of time before sorting both half-decks.

Sample 6. Seventy school students applied a less elaborate Grouping-and-Addition sorting procedure to the 55 CCQ(m) items. The Partitioning phase was omitted to ensure that participants knew what to do.

As well as sorting by similarity of meaning, the same participants Q-sorted the items into seven piles with eight items per pile. Their instructions were to use as target someone they knew well, of either sex, not necessarily someone they liked. An item was added to the CCQ(m) set to make the number divisible by seven, but it is ignored in the course of data analysis.

Analysis: Three 55-by-55 proximity matrices were derived for the CCQ(m) items. These were

D[Sm1], from the G, P, A data of Sample 5;

D[Sm2], from the G, P data of Sample 6; and

D[Qm], from the Q-sorts of Sample 6.

P -dimensional solutions were retained for each matrix (Sm1 _{p} , Sm2 _{p} and Qm _{p}), for $P = 4$ and 5 . We also combined D[Sm1] and D[Sm2] to obtain group solutions Gm _{p} . We used CANCELL comparisons to compare Sm1₅ and Sm2₅ against each other, and both against Gm₅. Those solutions were then compared against Qm₅. Finally, Gm₅ was compared against G₅.

Results

The outcomes of CANCELL comparisons for the CCQ(m) solutions were similar to those of Study 1. Four dimensions were replicated (with $p < 0.01$) across the external structure of Qm₅ and the internal structure of Gm₅. Subsequent analyses focus on four dimensions.

The first three dimensions of Gm₄ arrange the CCQ(m) items in an incomplete spherical shell, so that a reasonably clear impression can be conveyed by projecting the

negative- $D1$ and positive- $D1$ hemispheres separately into the D2/D3 plane, as shown in Figure 3. There is a close similarity to Figure 2, which depicts the comparable result for the complete item set. Figure 3 ignores the fourth dimension of the MDS solution – a fourth distinction made among items by informants – but this axis is mutually recognisable between G₅ and Gm₅, according to the CANCELL comparison. It is less important than the first three dimensions.

The subjective resemblance between G₄ and Gm₄ can be demonstrated more objectively using the same indices as in Study 1: the Procrustes statistic ($g_1 = 0.04$) and the inter-distance correlation ($r = 0.94$). These indices also confirm that there are no substantive differences between external and internal structure for the CCQ(m). Comparing Gm₄ with Qm₄ gives $g_1 = 0.12$ and $r = 0.83$. Finally, even less difference was found between the two sorting-data solutions, Sm1₄ and Sm2₄: $g_1 = 0.05$ and $r = 0.91$. It seems that the change in data-collection procedure between Samples 5 and 6 did not affect the perception of proximities.

The key result here is that removing 45 CCQ items has caused only minor shifts in the locations of the 55 remaining items. Neither the informants’ dissimilarity judgements, nor the analysis of those judgements, were sensitive to the context in which they were made.

Discussion

This report is not directly concerned with the Five-Factor model of trait psychology (FFM), or with the ‘Big-5’ factors themselves. Nevertheless, it is worth examining our results for relevant implications.

At issue is whether the five-factor framework is an observation about personality structure *per se*, or about personality perception. One view is that the factors are empirical realities, as undeniable as the “fact that there are seven continents on earth” (or only six, if the definition excludes Australia). They invite explanation but do not require justification. To invoke King Lear, the reason that there are no more than seven planets is “Because they are not eight.” It may be, then, the seeming complexity and range of human individuality is deceptive, since a five-dimensional coordinate or profile characterises anyone.

In this view, the observed convergence of external and internal structure is caused by the former affecting the latter. Perhaps when informants judge inter-item similarity (internal structure), they make these judgements by recalling and generalising from the actual trait co-occurrence they experienced while interacting with others (external structure). It would seem that the young-teenage subjects were not prevented from providing similarity-judgement data (codified as Gm₄) by their lesser experience of empirical trait co-occurrence, relative to

adults. Moreover, their judgements were almost identical with those of the adult subjects who provided most of the data for G_4 .

Alternatively, causality might operate in the converse direction. Internal structure will modulate the subjects' responses if they complete their peer descriptions with the help of a generic 'working model' of trait inter-relationships. Presented with a list of traits to be rated, in peer- or self-description, a subject may be able to respond immediately to some items. However, other items may be harder (perhaps the right conditions for observing the target have never arisen). Rather than think hard about the question, the subject might find it easier to interpolate: "This problematic item is somewhere in the vicinity of these two other traits, so I will give it the average of the ratings I have already given to them."

Note also that very few traits are marked by a single, unambiguous action or activity; they must be inferred from patterns of behavior. Thus similarity judgements can be expanded into the form "If the target's behavior can be explained by trait X (e.g. "Is considerate and thoughtful of other people"), how well could it be explained instead by trait Y (e.g. "Tries hard to please other people")? The semantic relationships among trait descriptors can be understood as descriptions of the limitations on what we can observe.

Of course there remains the question of where the internal structure comes from. We should consider the possibility that some categories of cognitive and affective response are universal, or at least widespread – a lowest common denominator of human experience. Osgood and his colleagues asked informants to make distinctions among phenomena within a range of semantic domains (e.g. Osgood, 1971). Consistently across languages and cultures, the results of examining these data with factor analysis could be brought under the aegis of a common framework of three dominant factors: Evaluation, Potency and Activity. These protean qualities can be conveyed by contrasts such as *good-bad* and *nice-awful* (Evaluation), *fast-slow* and *noisy-quiet* (Activity), *strong-weak* and *deep-shallow* (Potency).

In short, we are proposing that the Big-5 factors are grounded in the mechanisms by which we all form personality impressions. Thus the FFM is a model of instruments rather than of the targets observed through those instruments; a codification or elucidation of 'folk psychology'. While the measures it captures are highly salient in casual interaction – suitable, perhaps, for delineating the characters of television drama – much remains outside their scope and requires the observational opportunities of closer acquaintance. Dabady, Bell and Kihlstrom (1999) described the Big-5 factors as "Blind data questions".

Thus the FFM may not match the full complexity of personality. On the other hand, it may overstate the complexity of internal structure. We found that a spatial model with fewer than five dimensions could convey the main features of personality description at this broad-brushstroke level. Five or more factors can be embedded within a three- or four-dimensional model (as vectors, or directions of increasing value), though they will not be orthogonal.

Multivariate regression was used to summarise abstractions such as Self-esteem and Ego-Resilience as vectors within the geometrical model of internal structure. The components of each vector reduce that abstraction to its relative proportions of Evaluation, Activity, Potency and *D4*. Other 'empirical templates' have been published. Some studies characterised an empirical property, assessed across a population, in terms of its correlations with the values assigned to each CCQ item in descriptions of the same population (e.g. Externalising Behavior problems: Caspi *et al.*, Table 5). Other studies contrasted a target group who possess some property and a control group without it, and published the difference between mean scores for each CCQ item (e.g. Disruptive Disorder: Caspi *et al.*, 1992, Table 4). Thus one avenue for further inquiry into the dimensions of trait space would be to embed these templates and reduce them to vectors in the same way.

A further extension would be to calculate vectors for Q-sort descriptions of individuals, and summarise each target's observed behavior in terms of dimensional components. Given the association between the Evaluation axis and constructs such as optimal ego functioning and 'social competence', we would expect the *D1* component of an individual's vector to predict social performance, and school or career success. A further expectation is that these forms of performance should be linked with the individual's scores on the Big-5 factors, since these are also associated with *D1* (the association being weakest for FI, Surgency).

Conclusion

It is no surprise that the semantic similarity relationships within a set of trait descriptors, when analysed with multidimensional scaling, could be reduced to an armature of axes analogous to Osgood's 'EPA' factors. It made little difference whether the similarity judgements were made in the context of all 100 descriptors, or within a 55-item subset. What may be more interesting is that the same structure also appears to govern the practical use of these descriptors to describe targets, and predicts the pattern of associations with independent variables. In other words, the internal and external structures of the CCQ are the same. It remains to be seen whether the same convergence occurs for other inventories of sentence-length descriptors.

Though a three-dimensional spatial model does not exhaust these structures, it captures their main features. Indeed, the model can be presented in two dimensions, by treating separately the two ‘hemispheres’ (positive and negative values of *DI*, identified as ‘Evaluation’). For comparison, Green and Walkey (1980) applied MDS to the intercorrelations among the 57 items of Form A of the Eysenck Personality Inventory. A two-dimensional solution was sufficient, since the study in effect dealt with a single hemisphere (scores for some items were reversed to bring them into the positive poles of their respective factors). Another precedent is Maraun’s (1997) analysis of correlations among NEO-PI scales; again, these could be accommodated within two dimensions, because only the positive poles of the bipolar scales were represented by points, thereby confining them to a single hemisphere in the present model. We have already noted the associations between the Big-5 factors and the Evaluation axis *DI*, so that positive-pole items are confined to positive values of *DI*.

An earlier MDS exploration of a lexicon of 60 trait adjectives (Bimler & Kirkland, 2007) produced a spatial model of internal structure that was equally compatible with published factor analyses, and had three principal dimensions, conforming to the same schema of Evaluation, Activity, Potency. There is some divergence when items in the two models are given factor labels according to those published results. For instance, agreeableness and emotional stability, which are distinct in the CCQ (Figures 1 and 2), were poorly resolved in the single-word lexicon (Bimler & Kirkland, Figure 2). Rather than regard this as a real difference between the two spatial models, we ascribe it to a difference between alternative sets of five factors. That is, the factor models operationalised by Goldberg (1992) and John *et al.* (1994), respectively selecting adjectives and CCQ items as markers, were not identical.

Thus the two instruments cover much the same territory, vindicating the lexical hypothesis. However, the fourth dimension or distinction made among CCQ items (replicated in a 55-item subset) is not so easily assimilated to the *D4* found for adjectives (glossed in Bimler & Kirkland, 2007, as ‘emotionality’ or ‘abstract / concrete’). In both cases the fourth dimension was of minor importance. Conceivably the significance of such a dimension could be inflated by making it easier to attend to, by reducing the range of variation and thereby the salience of another dimension: for instance, by presenting informants with items from only the positive poles of Big-5 factors.

Within this territory, we have already noted the sparse distribution of points with g_{11} close to zero, representing CCQ items with no Evaluative content.

Points are particularly sparse at the extremes of the *D3* axis, i.e. the poles of positive and negative Potency. But ‘potency’ carries a connotation of potential rather than of manifest behavior: almost by definition it is hard to observe or operationalise, and the scarcity of examples is understandable.

With these exceptions, the CCQ samples the three-dimensional circumplex well. There is evidence that the pattern of perceived similarities alters little when items are removed or added, in the form of voids in Figure 2 where interstitial items were omitted between the clusters of factor markers. It follows that if the CCQ contained additional ‘blind spots’, they would also have appeared as voids in Figure 1. We are also encouraged to believe that an earlier spatial model for a trait lexicon (Bimler & Kirkland, 2007) was not seriously distorted by its paucity of interstitial items. That lexicon been selected to emphasise simple structure, so points were tightly clustered rather than forming a continuum in the MDS solution.

Behind the consensus about the importance of five factors in trait psychology, there is more than one five-factor model in contention. As noted above, the ‘Big-5’ factors emphasised in Figures 1 and 2 are not identical to the factors that have dominated studies of trait adjectives. The NEO-PI schema is different again. Johnson and Ostendorf (1993) have promoted a multidimensional circumplex structure (Hofstee *et al.*, 1992), because it is equally compatible with rival selections of factors, and can make explicit the differences between them (“One ring to bring them all, and in the darkness bind them”).

References

- Bimler, D.L. & Kirkland, J. (2007). Constructing personality maps, mapping personality constructs: Multidimensional scaling recovers the ‘Big Five’ factors from internal and external structure. *Spanish Journal of Psychology*, 10, 68-81.
- Block, J. (1961). *The Q-sort method in personality assessment and psychological research*. Springfield, IL: Charles C Thomas (reprinted 1978, Palo Alto, CA: Consulting Psychologists Press).
- Block, J. & Block, J.H. (1980). *The California Child Q-set*. Palo Alto, CA: Consulting Psychologists Press.
- Briggs, S.R. (1992). Assessing the five-factor model of personality description. *Journal of Personality*, 60, 253-293.
- Caspi, A., Block, J., Block, J.H., Klopp, B., Lynam, D., Moffitt, T.E. & Stouthamer-Loeber, M. (1992). A “common-language” version of the California Child Q-set for personality assessment. *Psychological Assessment*, 4, 512-523.
- Cattell, R.B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38, 476-507.

- Church, A.T. & Katigbak, M.S. (1989). Internal, external, and self-report structure of personality in a non-Western culture: An investigation of cross-language and cross-cultural generalizability. *Journal of Personality and Social Psychology, 57*, 857-872.
- Dabady, M., Bell, M. & Kihlstrom, J. F. (1999). Person memory: organization of behaviors by traits. *Journal of Research in Personality, 33*, 369-377.
- Fitzgerald, H.E., Zucker, R.A. & Yang, H.-Y. (1995). Developmental systems theory and alcoholism: Analyzing patterns of variation in high-risk families. *Journal of Addictive Behaviors, 9*, 8-22.
- Goldberg, L.R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26-42.
- Goodhill, G.J., Simmen, M.W. & Willshaw, D.J. (1995). An evaluation of the use of multidimensional scaling for understanding brain connectivity. *Philosophical Transactions of the Royal Society of London B, 348*, 265-280.
- Green, D.E. & Walkey, F.H. (1980). A nonmetric analysis of Eysenck's Personality Inventory. *Multivariate Behavioral Analysis, 15*, 157-163.
- Gurtman, M.B. & Pincus, A.L. (2000). Interpersonal Adjective Scales: Confirmation of circumplex structure from multiple perspectives. *Personality and Social Psychology Bulletin, 26*, 374-384.
- Guttman, L. (1966). Order analysis of correlation matrices. In *Handbook of Multivariate Experimental Psychology* (ed. R.B. Cattell), 438-458. Chicago: Rand McNally.
- Hofstee, W.K.B., de Raad, B. & Goldberg, L.R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology, 63*, 146-163.
- John, O.P., Caspi, A., Robins, R.W., Moffitt, T.E. & Stouthamer-Loeber, M. (1994). The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65*, 160-178.
- Johnson, J.A. & Ostendorf, F. (1993). Clarification of the Five-Factor Model with the Abridged Big Five Dimensional Circumplex. *Journal of Personality and Social Psychology, 65*, 563-576.
- Kogan, N. & Block, J. (1991). Independence from early childhood through adolescence: Personality and socialization aspects. In *Field Dependence-Independence: Cognitive style across the life span* (eds. S. Wapner & J. Demick), pp. 177-207. Hillsdale, NJ: Lawrence Erlbaum.
- Lanning, K. (1994). Dimensionality of observer ratings on the California Adult Q-set. *Journal of Personality and Social Psychology, 67*, 151-160.
- Lorr, M. (1978). The structure of the California Q-set. *Multivariate Behavioral Research, 13*, 387-393.
- Maraun, M.D. (1997). Appearance and reality: Is the Big Five the structure of trait descriptors? *Personality and Individual Differences, 22*, 629-647.
- McCrae, R.R. & Costa, P.T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.
- McCrae, R.R. & Costa, P.T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality, 57*, 17-40.
- McCrae, R.R., Costa, P.T. & Busch, C.M. (1986). Evaluating comprehensiveness in personality systems: The California Q-set and the five-factor model. *Journal of Personality, 54*, 430-446.
- Osgood, C.E. (1971). Explorations in semantic space: A personal diary. *Journal of Social Issues, 27*, 5-64.
- Peabody, D. & Goldberg, L.R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology, 57*, 552-567.
- Schlesinger, I.M. & Guttman, L. (1969). Smallest space analysis of intelligence and achievement tests. *Psychological Bulletin, 71*, 95-100.
- Sherman, R.C. (1972). Individual differences in perceived trait relationships as a function of dimensional salience. *Multivariate Behavioral Research, 7*, 109-129.
- Sneed, C.D., McCrae, R. R. & Funder, D.C. (1998). Lay conceptions of the five-factor model and its indicators. *Personality and Social Psychology Bulletin, 24*, 115-126.
- van Lieshout, C.F.M. & Haselager, G.J.T. (1994). The Big Five personality factors in Q-sort descriptions of children and adolescents. In: C.F. Halverson, G.A. Kohnstamm and R.P. Martin (eds.), *The Developing Structure of Temperament and Personality* (pp. 293-318). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Vaughn, B.E. & Martino, D.G. (1988). Q-sort correlates of visual regard in groups of young pre-school children. *Developmental Psychology, 24*, 589-594.
- Waters, E., Garber, J., Gornal, M. & Vaughn, B.E. (1983). Q-sort correlates of visual regard among preschool peers: Validation of a behavioral index of social competence. *Developmental Psychology, 19*, 550-560.
- Waters, E., Noyes, D.M., Vaughn, B.E. & Ricks, M. (1985). Q-sort definitions of social competence and self-esteem: Discriminant validity of related constructs in theory and data. *Developmental Psychology, 21*, 508-522.
- Wiggins, J.S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.
- Zucker, R.A., Fitzgerald, H.E., Refior, S.K., Puttler, L.I., Pallas, D.M. & Ellis, D.A. (2000). *The clinical and social ecology of childhood for children of alcoholics: Description of a study and implications for a differentiated social policy*. In: H.E. Fitzgerald, B.M. Lester, & B.S. Zuckerman (eds.), *Children of addiction: Research, health and policy issues* (pp. 109-141). New York: Routledge/Falmer.

Received August 28, 2008

Revision received May 11, 2009

Accepted July 30, 2009