Universality of the Five-Factor Model of Personality

Velichko H. Fetvadjiev
University of Pretoria, South Africa

Fons J. R. van de Vijver
Tilburg University, The Netherlands, North-West University, South Africa, and University of Queensland, Australia


Author Note

Velichko H. Fetvadjiev
Department of Human Resource Management, University of Pretoria. Email: velichko.fetvadjiev@gmail.com. Until 2014, Velichko H. Fetvadjiev was publishing by the family name of Valchev.

Fons J. R. van de Vijver
Department of Culture Studies, Tilburg University, The Netherlands. Email: fons.vandevijver@uvt.nl. North-West University, South Africa, and University of Queensland, Australia.
Abstract

The Five-Factor Model (FFM) is a language-based model of the structure of personality. Evidence from 3 lines of research on the FFM’s universality across cultures is reviewed: (a) questionnaire studies that sought to replicate the FFM across cultures; (b) psycholexical studies that examined implicit models of personality in different languages using lexica; (c) indigenous and emic–etic studies that examined personality structures in specific cultures using a variety of methods, and subsequently assessed the overlap between these models and the FFM. Support for the universality of the FFM is strongest when FFM-based questionnaires are used and tends to be stronger in Western than in non-Western cultures. Four factors are usually well replicated (Neuroticism, Extraversion, Agreeableness, and Conscientiousness), whereas the fifth factor, Openness, is not always found. From a psycholexical perspective, the replicability of 3 broad dimensions is best established: Extraversion, Agreeableness, and Conscientiousness. Emic–etic studies tend to converge on a structure similar to the FFM although sometimes expanding it in the domain of social-relational functioning. The non-replication of the FFM cannot be easily attributed to a common pattern of causes; data quality, availability of abstract terms in the language, and cultural salience of the concepts are often involved.

Keywords: personality, Big Five, FFM, culture, universality, country profiles, psycholexical study, indigenous psychology, emic–etic approach, cross-cultural psychology, cultural psychology
The Five-Factor Model (FFM) of personality is a model of the structure of personality firmly based in language, according to which human personality can be described using five broad dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. Each of these dimensions subsumes a number of lower level traits on both the positive and the negative pole: for example, *talkative* and *reserved* in Extraversion, *friendly* and *harsh* in Agreeableness, *orderly* and *messy* in Conscientiousness, *calm* and *fearful* in Neuroticism, and *perceptive* and *ignorant* in Openness (Goldberg, 1990). The terms *Five-Factor Model* and *Big Five* are often used interchangeably, although in the narrow sense, the former refers to a stricter version of the model with 30 definite lower level facets (Costa & McCrae, 1992), whereas the latter encompasses a broader array of models with different lower level components; we follow the habit of using the two terms as synonymous.

The FFM was initially developed to provide a parsimonious account of variation in personality descriptions in the English language. The FFM dates back to work by Allport and Odbert (1936) and Stagner (1937) in the 1930s, revived in the 1960s by Norman (1963) and Tupes and Christal (1961), and widely established in the last decades of the 20th century by various researchers, notably Goldberg (1990) and McCrae and Costa (1997). This renewed interest spurred cross-cultural studies of the FFM, turning the attention toward the cross-cultural replicability of the model. If the FFM were found to replicate across cultures and languages, this would provide support for its validity as a panhuman model of personality and would allow it to be used as a basis for theorizing and empirical research transcending cultural and regional boundaries. In parallel developments, research on animal personality has identified personality components that could be interpreted in terms of (subsets of) the
FFM in animals, whereby most overlap was found in humans’ closest relatives, chimpanzees (Gosling & John, 1999).

To what extent is there a single, species-wide model of personality in humans? The positions taken on this question vary. Historically, the discussion goes back to the emic–etic distinction, introduced by Pike (1967). “Etic” refers to research that studies cross-cultural differences from a comparative perspective (studying each individual culture less profoundly), whereas “emic” refers to research that examines a specific culture in more detail with no (or only a secondary) cross-cultural focus. An etic, universalistic perspective in personality can be found in five-factor theory (McCrae & Costa, 2008), where the dimensions of personality are assumed to be biologically based and universal, whereas specific manifestations of the model’s components are shaped by culture. From a more relativistic perspective, personality can only be understood in cultural and historical context (Piekkola, 2011). There are methodological approaches, such as the psycholexical approach (discussed in a following section), which have no a priori position on the universality of the FFM or other models, but provide important tools for addressing the topic. Finally, there are theoretical and methodological approaches, labeled emic–etic (Cheung, Van de Vijver, & Leong, 2011), that seek to combine universalistic and relativistic approaches. Emic–etic approaches seek to integrate evidence on more universal and more culture-specific elements in personality models.

We review the empirical evidence on the FFM accumulated from three lines of research: (a) questionnaire studies that sought to replicate the FFM across cultures; (b) psycholexical studies that sought to establish implicit models of personality in different languages using lexica; (c) indigenous and emic–etic studies that sought to establish personality structures in specific cultures using a variety of methods, and
subsequently assessed the overlap between these models and the FFM. In the first two
two sections, we start from the evidence in support of FFM universality and progress
toward divergent evidence, while also commenting on methodological limitations and
shortcomings. In the final section, we present findings on indigenous concepts and
their degree of overlap with the FFM in several non-Western cultures.

**Questionnaire Studies**

In the questionnaire study design, participants respond to a list of questions
about their typical characteristics, behaviors, feelings, and thoughts. Items include
abstract trait terms (e.g., “I am talkative”) as well as behavioral terms (e.g., “I make
friends easily”). Responses are factor-analyzed, and the resulting structure is
compared with an existing factor structure, typically derived from U.S. data.
Statistical criteria are used to evaluate agreement by using congruence coefficients
that indicate the degree of equivalence between the two structures (interpreted as
supportive evidence for the identity of factors across the cultures studied).
Congruence coefficients of .95 and higher indicate equivalent structures (Lorenzo-
Seva & Ten Berge, 2006; Van de Vijver & Leung, 1997), although values of .85 or
higher are sometimes also taken as evidence of similarity. A similar procedure can be
used to compare individual-level to country-level factor structures.

The most popular inventory for measuring the FFM is the Revised NEO
Personality Inventory (NEO-PI-R; Costa & McCrae, 1992), which measures 5 broad
domains and 30 lower level facets, 6 per domain (i.e., the FFM in the narrow sense,
mentioned in the Introduction). Another popular instrument is the Big Five Inventory
(BFI; John, Donahue, & Kentle, 1991). Both instruments have been translated into a
large number of languages from different families, using back-translation procedures.
The self-report form of the NEO-PI-R has been administered in 36 cultures (McCrae,
2002), and the peer-report in 50 cultures (McCrae, Terracciano, & 78 Members of the Personality Profiles of Cultures Project, 2005a). The BFI has been administered in 56 cultures (Schmitt, Allik, McCrae, Benet-Martínez, et al., 2007). All these studies tended to find fair to very good replicability of the factors of the FFM in individual cultures and for the total set of cultures compared with the U.S. structure. The general replicability between the individual- and culture-level factor structure, together with evidence for the external validity of culture-level scores (e.g., their associations with cultural values and economic indices), has enabled researchers to construct “personality profiles” of cultures (McCrae, Terracciano, & 79 Members of the Personality Profiles of Cultures Project, 2005b) and to explore the geographical distribution of Big Five traits (Schmitt et al., 2007). For example, Schmitt et al. (2007) found that average levels of Extraversion were lower in East Asia than in other regions. Finally, broad support for the replicability of the FFM has also been found in studies with less extensive culture coverage, both using other instruments designed to measure the FFM (Caprara, Barbaranelli, Bermúdez, Maslach, & Ruch, 2000; De Fruyt et al., 2009; Hendriks et al., 2003) and using instruments that had not been initially devised as a measure of the FFM, including nonverbal instruments (Paunonen et al., 1996).

Several limitations on the support for the universality of the FFM from questionnaire studies require attention. First, although the five factors generally tend to replicate, not all replications pass accepted thresholds for construct equivalence (Lorenzo-Seva & Ten Berge, 2006; Van de Vijver & Leung, 1997). More discrepancies tend to be found at lower levels of aggregation (i.e., for individual countries and country regions) than for global solutions. In McCrae et al. (2005a), there were a number of non-Western, especially African, countries with congruence
coefficients below .80, especially for Openness. The authors attributed the finding to shortcomings in data quality rather than insufficiency of the model for these countries. In Schmitt et al. (2007), several coefficients especially in Africa and South and Southeast Asia were below .90, with the lowest one of .84 for Openness in Africa. Second, although the culture-level five factors map fairly closely on the individual-level factors, the two structures are not identical. On culture level, Extraversion was co-defined by a number of facets from other factors (McCrae et al., 2005b), and the correlations between the NEO-PI-R and BFI factors were less supportive of the convergent and discriminant validity of the factors than on individual level (Schmitt et al., 2007). The implication is that interpretations of individual differences cannot be transferred simply to interpretations of country mean differences. Third, although the structural equivalence of the five factors on individual level has broadly been established, there is no comparable support for measurement equivalence and there is some evidence for inequivalence (Church et al., 2011). The implication is that direct numerical comparisons of trait scores across cultures are not yet warranted: Individual differences in the five factors do not have the same meaning across countries. Fourth, the questionnaire study design has the inherent limitation that it can only be used in a straightforward manner in literate populations. Most of the studies have been conducted on student samples. In one of the rare studies on preliterate societies, Gurven, Von Rueden, Massenkoff, Kaplan, and Lero Vie (2013) failed to replicate the FFM using the BFI. The authors interpreted their finding as evidence that the FFM may not be applicable to the Tsimane ethnic group in the Bolivian Amazon and likely other small-scale, preliterate societies; a more conservative interpretation would involve issues of method bias and inadequacy of questionnaire adaptation. Finally, an inherent limitation of questionnaires is that they tend to impose structure on the data.
to a certain extent (Church, 2000). Support of cross-cultural replicability has been obtained also for other models, measured with respective questionnaires, such as the Eysenck Personality Questionnaire (EPQ, Barrett, Petrides, Eysenck, Eysenck, 1998). While questionnaires offer a powerful tool for replicating factor structures, they are less efficient as tools for discovering factors implicit in different cultures and languages. Such questionnaires cannot address the question whether the items are the best possible rendering of personality in that culture and whether culture-specific aspects were not left out.

**Psycholexical Studies**

Research in the psycholexical (or lexical) tradition is based on the assumption that the most important features of personality become encoded as single terms in language ("psycholexical hypothesis"; Goldberg, 1981). Single personality-descriptive terms can be derived from lexica and subsequently reduced through content and factor analysis to arrive at a few broad dimensions of personality. A lexical study begins with an exhaustive extraction of personality terms from current dictionaries. In the stricter and more common version of the approach, only terms referring to internal psychological characteristics are selected, whereas terms that are purely evaluative, refer to physical appearance, beliefs, temporary states, or do not describe personality are excluded (Saucier, 2008). The list is further shortened by selecting the terms that appear most frequently in natural language usage. Participants are presented with the selected trait list and asked to assess to what extent each term is applicable to them (or to a target person they know well). The scores of the participants are factor-analyzed, resulting in models with a few dimensions that account for the covariation of trait terms. Because of the comprehensive selection of trait terms, this approach is less restrictive than the questionnaire approach in the
dimensions it can identify and is thought to be more representative of the implicit structure of personality in each language.

The Big Five model was first identified in the outcomes of lexical research in English (Goldberg, 1990). Lexical studies have since been conducted in about fifteen other, Indo-European as well as non-Indo-European languages. Direct empirical comparisons and conceptual overviews of the lexical data across languages have not converged on a single model replicable across languages (De Raad et al., 2010; Saucier, 2009). Although factors similar to those of the Big Five have been identified across languages, the Big Five as a coherent structure defined by common elements has only been replicated in the Germanic languages (Saucier & Goldberg, 2001). In an analysis on data from 12 languages obtained according to the common premises (e.g., full rather than random extraction of terms, exclusion of purely evaluative terms), with explicit definition of common terms defining factors, and direct psychometric assessment of structure equivalence (with a liberal cutoff value for congruence coefficients at .80), De Raad and colleagues (2010) found that only three factors were fully replicable across languages: Extraversion, Agreeableness, and Conscientiousness. These factors were more broadly defined than the corresponding Big Five factors because they included terms from the other factors. These three factors tend to be the three largest factors in lexical studies. There is less agreement across languages on factors beyond these three.

Some limitations of the psycholexical approach are effected from the opposing aims of arriving at a common model for implicit personality concepts across languages while providing a comprehensive representation of these concepts. The first aim drives research toward the identification of fewer personality factors, whereas the second drives research toward more factors (De Raad et al., 2010). Comparability of
factors across languages is challenged by procedural differences between studies with respect to the selection of data sets and of terms to be analyzed, item translation, and decisions regarding factor extraction and rotation. In addition, there is the more important limitation that different sets of items provide the input for the analyses in different languages. The factors are usually defined by a limited number of highly loading items, called *markers*. These markers tend not to overlap across languages, and decisions have to be made as to which markers to use for factor identification and subsequent comparisons. These decisions can be biased by prior expectations, for example, of retrieving the Big Five factors. A classic example is that of Assertiveness terms. In the English Big Five structure, Assertiveness is a facet of the first factor, Extraversion. When, in other languages, Emotional Stability fails to emerge as a fourth factor, Assertiveness terms loading on the fourth factor can be taken to indicate Emotional Stability because this is the expected factor. Such interpretations amount to circular reasoning in assuming what has to be shown (De Raad & Peabody, 2005).

A different set of limitations apply to the conceptual premises of the psycholexical approach. Although it is reasonable to expect that trait terms are used universally, it has not been established that they are equally frequent across languages and perceived as equally informative of personality across cultures. There is evidence that abstract trait terms are used more often in personality descriptions in Western groups, whereas specific behavioral descriptions are more preferred in non-Western groups (Del Prado et al., 2007; Valchev, Van de Vijver, Nel, Rothmann, & Meiring, 2013). It may also be the case that some personality domains are less saturated with single abstract trait terms, which could explain the weak replicability of the smaller factors, Neuroticism and Openness, in lexical studies as opposed to questionnaire studies. In more practical terms, the psycholexical approach is not applicable to
languages where there are wide variations between a common written language and spoken dialects, as in Arabic, and to languages that have no developed lexicography.

**Indigenous and Emic–Etic Studies**

In contrast to the etic studies that seek to replicate personality structure across cultures, such as the NEO-PI-R replicability studies, indigenous or emic studies focus on the identification of personality concepts perceived as relevant in specific, typically non-Western cultures. Emic studies employ an array of methods including psycholexical methods, ethnographic interviews, and reviews of folklore and literature for personality characteristics. The strength of this kind of research is in the comprehensive representation of the implicit personality models, not bound by questionnaire item selections or exclusive reliance on lexical records.

There is some work on individual, purely indigenous (emic) personality concepts, such as the East-Asian, religion-based selfless self (Ho, 1995) and the Japanese Amae (Doi, 1973). However, this work has never developed into full-fledged personality models. When the personality concepts derived in indigenous research are compared with those from other cultures or with supposed universal models, more confident conclusions on the generalizability and universality of personality concepts can be drawn. Some of the major findings from indigenous research in different parts of the world are presented below.

In the Philippines, indigenous research has identified a number of implicit personality dimensions represented in different models (Guanzon-Lapeña, Church, Carlota, Katigbak, 1998). Katigbak and colleagues (Katigbak, Church, & Akamine, 1996; Katigbak, Church, Guanzon-Lapeña, Carlota, & Del Pilar, 2002) examined the indigenous dimensions jointly with the FFM and found that most indigenous concepts could be subsumed in the FFM. There were only a few concepts that were not
empirically encompassed by the FFM, such as social curiosity, risk-taking, and religiosity (Katigbak et al., 2002). Importantly, the indigenous concepts offered only modest incremental criterion validity above the FFM.

In Mexico, Ortiz et al. (2007) reviewed previous literature on indigenous concepts and examined the relations of a large number of indigenous measures with the FFM. They found that most of the Mexican concepts could be well subsumed within the FFM. Concepts of warmth and affection appeared to be more salient in the indigenous measures, whereas Openness could not be identified in the indigenous measures but only in the imported FFM measures. Overall, both in the Philippines and in Mexico, researchers were led to conclude that cultural differences may not be expressed in the clear identification of dimensions beyond the Big Five, but rather in the salience and composition of elements that can all be subsumed in the Big Five.

In China, several independent lines of research have explored the indigenous structure of personality (reviewed in Yang, 2006). The best established questionnaire that resulted from one of these lines of research is the Cross-Cultural (initially Chinese) Personality Assessment Inventory (CPAI-2, Cheung, Cheung, Leung, Ward, & Leong, 2003). Developed with input from interviews and reviews of wide-ranging literature sources, this inventory identified one dimension that could not be subsumed within the FFM, Interpersonal Relatedness, which involves aspects of functioning in social context and adherence to tradition. Subsequent studies have replicated the Interpersonal Relatedness factor in non-Chinese populations, including the U.S. (Lin, & Church, 2004), suggesting that although Interpersonal Relatedness is most salient in China, it is not unique to China but is recognizable in other cultures. Openness, on the other hand, has not been identified in China using indigenous measures but only using imported measures (Cheung et al., 2008). So, although single concepts referring
to Openness are used in the Chinese context, they are not implicitly recognized as constituting a coherent personality dimension. This pattern is opposite to that of Interpersonal Relatedness. Finally, it is noteworthy that the Interpersonal Relatedness dimension has been found to offer incremental validity in the prediction of culturally relevant outcomes in China (Zhang & Bond, 1998).

In South Africa, a culturally heterogeneous country with 11 official languages, recent indigenous research has identified nine clusters of personality concepts derived from free personality descriptions (Nel et al., 2012). All of the Big Five factors were recognizable within the nine clusters, but the nine clusters included a large number of concepts dealing with interpersonal relations. This finding is consistent with findings from lexical and other indigenous studies, where Agreeableness-related terms normally form the largest factor. The quantitative validation of the conceptual model is currently ongoing, but there are some initial indications that concepts related to social-relational functioning in a fairly norm-regulated context are salient in South Africa, yet not well covered by the FFM.

Conclusion

Support for the universality of the FFM is strongest when FFM-based questionnaires are used and tends to be stronger in Western than in non-Western cultures. Four out of the five factors are usually well replicated (Neuroticism, Extraversion, Agreeableness, and Conscientiousness), whereas the fifth factor, Openness, is not always found. From a psycholexical perspective, the replicability of three broad implicit dimensions is best established: Extraversion, Agreeableness, and Conscientiousness. The non-replication of the FFM cannot be easily attributed to a common pattern of causes; data quality, availability of abstract terms in the language, and cultural salience of the concepts are often involved. Despite some strong
proposals (e.g., Buss, 1991; McCrae & Costa, 2008), there is no commonly accepted theory of the Big Five or FFM that would account for the universal occurrence of the five factors or their causal role in behavior. In fact, the role of the Big Five factors as causal entities is periodically challenged (e.g., Cramer et al., 2012). The lack of such a commonly accepted theory can be perceived as contributing to the dangers of circular reasoning (De Raad & Peabody, 2005). Nonetheless, research on the universality of the FFM has produced invaluable data and rich insights in the patterns of similarities and differences in personality across cultures.
References


Cross-References

Cross-Cultural Psychology, 24011
Cross-Cultural Research Methods, 44014
Cultural Psychology, 24053
Cultural Relativism, Anthropology of, 12047
Indigenous Psychology, 24067
Five Factor Model of Personality, Assessment of, 25061
Five Factor Model of Personality, Facets of, 25063
Theories of the Big Five Factor Model, 25066