Ubiquitous Personality-Trait Concepts in 13 Diverse and Isolated Languages: A Cluster–Classification Approach

JOSHUA K. WOOD[1]*, MICHAEL GURVEN[2] and LEWIS R. GOLDBERG[3]

[1]*School of Psychology, Deakin University, Geelong, Australia*
[2]*Department of Anthropology, University of California at Santa Barbara, USA*
[3]*Oregon Research Institute, USA*

Abstract: There is longstanding interest in the generalizability of personality across diverse cultures. To investigate the generalizability of personality concepts, we examined the English translations of individual-difference entries from the dictionaries of 12 small-scale societies previously studied for ubiquity of individual differences plus the dictionary of an additional society not previously studied in this manner. These 13 societies are highly diverse in geographical location, culture, and language family; their languages developed in isolation from modern-world languages. The goal of our exploratory research was to discover ubiquitous personality concepts in these 13 independent societies and their languages, providing a window into personality concepts across a broad range of cultures and languages. This study used clusters of empirically related terms (e.g. brave, courageous, and daring), based on a taxonomy of English-language personality concepts that consisted of 100 personality-trait clusters. English-language definitions of dictionary entries from the 13 languages were matched to the meanings of the synonym clusters. The cluster–classification method uncovered nine ubiquitous personality concepts, plus six that were present in at least 12 of the 13 languages. The nine ubiquitous personality concepts include some not previously identified and suggest a core of possibly universal concepts. © 2020 European Association of Personality Psychology

Key words: personality; Big Five; lexical study; cross-cultural; ubiquitous

> It is one thing to be the servant of language and another to make use of its hints. (Klages, 1926/1932, p. 41)

Suppose there are some personality-trait concepts that are used by members of every society in the world, no matter how isolated. The universality of these concepts across such disparate, independently derived origins might indicate that these personality differences are intrinsic to all humans and so universally relevant that human beings in all types of societies notice them.

The *lexical hypothesis* is discussed in detail by Saucier and Goldberg (1996, 2001), but its simplest form, first expressed by Klages (1926/1932), is that human attributes that are important to people in a society will become part of that society's language. Social relevance is believed to determine 'importance', where explicit recognition of other society members' traits enables better decisions on how to interact with them (Srivastava, 2010). This hypothesis and its second part, that the more important attributes are likely to become encoded in language as *single words*, have been expressed in various forms by Cattell (1943), Norman (1963, 1967), and Goldberg (1981, 1982).

*Correspondence to: Joshua K. Wood, School of Psychology, Deakin University, Geelong, Australia
E-mail: woodjos@deakin.edu.au

Saucier and Goldberg (1996) extended the simple form of the lexical hypothesis by proposing that an attribute's level of importance is associated with its degree of representation in language. This hypothesis applies across languages, whereby the more important human attributes will have accompanying terms in more languages (Goldberg, 1981). However, it also applies within languages, with Saucier and Goldberg (1996) expressing that the more important attributes will be (i) crowded more closely with a cluster of synonyms, (ii) more frequently used, and (iii) more highly correlated with other terms and so will be central to the definition of a semantic factor.

Only some parts of the lexical hypothesis have been empirically tested, with mixed results. For example, the expectation that important terms would have more synonyms was not supported in a study by Wood (2015), where US participants rated the importance of commonly used terms. However, Wood found that socially important terms were more highly correlated and more likely to load highly on principal component factors. The expectation that important terms would be used more frequently was not supported by Wood but was supported by Leising, Scharloth, Lohse, and Wood (2014) with German participants and German terms.

The idea that the most important individual differences in people's lives become encoded as single words can be used as a justification for distilling individual-difference terms from the dictionaries of a single language and creating a wide-ranging classification of human attributes. Such a task

was accomplished by Allport and Odbert (1936), who extracted nearly 18 000 such terms from a comprehensive English-language dictionary. Subsets of these terms have since been used as a starting point for many psycholexical studies. For a history of the lexical approach, see John, Angleitner, and Ostendorf (1988).

However, given the across-language aspect of the hypothesis—that the more important attributes will have a term in more languages and the most important will have a word in nearly all languages (Saucier & Goldberg, 1996, 2001)—the lexical hypothesis is also a useful starting point for investigating the prevalence of personality concepts across different cultures. Being found in nearly all languages provides reasonable evidence for the cross-cultural importance of attribute terms (Saucier & Goldberg, 2001). Furthermore, terms that are widespread across languages may also be useful for identifying universals (Saucier & Goldberg, 1996).

There are multiple methods for determining the ubiquity of personality concepts across cultures, and each approach has its own methodological strengths and weaknesses. One method is to search for English words that are common to a set of indigenous dictionaries (e.g. Saucier, Thalmayer, & Bel-Bahar, 2014). However, if one were to search for ubiquitous person-descriptive concepts by counting the number of languages that included a particular term (in the English translations of the indigenous language terms), perhaps only half of the world's languages would include a term that had been translated as *intelligent*, while the other half might include a term that had been translated as *smart*. One might conclude that neither of these two terms was ubiquitous, much less universal, among the languages of the world. On the other hand, it would be justifiable to conclude that the person-descriptive *concept* (i.e. intelligent/smart) was ubiquitous if it turned out that all languages included a term that had been translated into one of the two words in that synonym pair.

The problem with using English to translate terms from other languages is that the enormous contemporary English lexicon includes far more terms than any other language spoken on the planet. This is due in part to the Norman invasion of England in the 11th century AD and the subsequent influence of German and Romance language families. In English, the abundance of synonyms complicates translation, as concepts can be translated into one or more terms derived from various modern-world languages as well as from Latin and ancient Greek.

Personality traits are concepts, and each personality concept may have one dominant high-frequency term and/or a number of lexicographically distinct exemplar terms, all of which can be considered a set of synonyms associated with that personality trait. Counting person-descriptive concepts, represented by *clusters of synonym terms*, instead of one-term descriptors, would, in many cases, be a more inclusive way of measuring the ubiquity of a person-descriptive concept. Therefore, we need a reliable and defensible way to group together terms that are empirically similar enough to be considered part of the same person-descriptive concept. A pre-established taxonomy of English-language individual-difference terms would be especially valuable—one with quasi-synonyms grouped together into traits based on

inter-item correlations and the use of synonym finders. Fortunately, such a taxonomic framework exists (Goldberg, 1990), and we use this catalogue to answer the following question: *What are the most ubiquitous personality concepts in the languages of the world?*

PREVIOUS UBIQUITY RESEARCH AND SOURCE MATERIAL FOR THE CURRENT STUDY

In a groundbreaking article, Saucier et al. (2014) used carefully chosen indigenous-to-English dictionaries to establish the relative ubiquity of individual-difference terms across 12 diverse languages: Afar (native to Djibouti, Eritrea, and Ethiopia), Enga (Papua New Guinea), Fijian (Fiji), Hmong (China, Laos, Thailand, and Vietnam), Hopi (North America), Inuktitut (Canada), Khoekhoe (Botswana, Namibia, and South Africa), Kuna (also known as Cuna and most recently as Guna; Colombia and Panama), Mara Chin (India and Myanmar, formerly Burma), Maasai (also known as Maa; Kenya and Tanzania), Supyire (Ivory Coast and Mali), and Wik-Mungkan (Australia).

As Saucier et al. (2014) noted, universality of human-attribute concepts requires examination of all cultures and languages, but it is currently only feasible to study a small portion of these languages in lexical studies. However, to maximize the representativeness and independence of the languages examined, Saucier et al. selected 12 languages from cultures that were isolated from each other, geographically distant (multiple continents), diverse in cultural features, and from a variety of language families. They focused on 'traditional' societies that, at the time when their respective dictionaries were written, were relatively independent from national and global cultural influences. That these societies and their languages developed *independently* means that any commonality in personality terms is likely to be generalizable to languages and peoples from a broad range of societies. Using English as the common translation language, Saucier et al. examined the dictionaries of these 12 languages, with each dictionary containing English translations of native terms. Although these dictionaries were selected as the best ones available, they appeared to differ in their inclusiveness of terms. Therefore, the absence of an entry in a dictionary does not definitively mean that a concept cannot be expressed in that language. Saucier and colleagues extracted 16 857 human-attribute terms from these 12 languages. Ubiquity, in its most stringent form, was considered to occur if all 12 languages had an exact match for one particular English term, or an alternate based on the same root, used within the English-translated definition of at least one word. To determine ubiquity, they first split all English translations from the 12 languages into their individual words (e.g. the Fijian word *somisisi*, translated as 'curious, inquisitive, impertinent' in the Fijian-to-English dictionary, was split into three individual words: "curious", "inquisitive", and "impertinent"), and then they counted the number of languages in which each word appeared. This method was chosen in order to maximize the objectivity and replicability of the process.

Furthermore, Saucier et al. (2014) cast a wide net by applying a systematic criterion of terms that could differentiate people. As such, their terms included temporary states (e.g. sad), physical characteristics (e.g. tall), social status (e.g. wealthy and married), evaluative concepts (e.g. good and useless), and behavioural propensities (e.g. clumsy and quick), among others. This broad approach avoided the careless exclusion of any individual-difference terms that might be judged to not be personality attributes or to be irrelevant.

Saucier et al. (2014) discovered 28 single words present in all 12 languages. Of these words, 10 would probably be considered conventional personality traits (bad, good, useless, disobedient, and stupid) or emotional states that might also be used to characterize people's personality traits (afraid, angry, ashamed, jealous, and surprised), five were related to physical attributes (e.g. big and beautiful), seven were related to health or physical potency (e.g. well, strong, sick, weak, tired, and blind), two were related to age (old and young), and four were potentially relevant, but it was unclear whether the terms applied to humans in all languages (e.g. cold and dirty). If *strong* and *weak* [which were included in Goldberg's, 1990 566 trait adjectives (566-TDA)] and *tired* were included, then the total number of terms considered conventional personality traits or states could be extended to 13.

Using a less stringent criterion of near ubiquity, where a term is present in 11 out of the 12 languages, yielded an additional 41 terms, of which 14 (including *sleepy* but not including *drunk*) could be considered conventional personality traits or states. These 41 terms included four fundamental motivational state terms—hunger, thirst, pain, and pleasure—that would perhaps be expected to occur in all of the languages, suggesting that the dictionaries varied in their inclusiveness.

For nearly half of the near ubiquitous terms, it was the Kuna dictionary missing the terms. The Kuna dictionary, as well as the Enga dictionary, included the fewest human-attribute entries, and these two dictionaries were the source of the most nonubiquity. While the small number of human-attribute entries in the Kuna dictionary (306, equaling 5% of the entries) suggests that it may not be a very inclusive dictionary of the Kuna language, it is also possible that Kuna simply has a small lexicon of personality concepts.

In order to make a direct comparison between Kuna and another society from South America, we added a dictionary from a 13th language, Tsimane. The Tsimane language is used by a forager–horticulturalist population from central lowland Bolivia (Gurven, von Rueden, Massenkoff, Kaplan, & Lero Vie, 2013). The Tsimane location and language are not directly related to that of the Kuna people, who are located in Panama and Colombia. The Tsimane population numbers approximately 16 000 in total but they live in smaller, relatively egalitarian communities of 30 to 500. They are closer to being foragers or semisedentary farmers than the other societies from the Americas studied by Saucier et al. (2014), whereas Kuna is an agricultural society that includes keeping livestock, travelling to hunt and fish, cloth manufacturing, and long-distance trading. Table 1, adapted from Saucier et al., lists the 13 languages and 15 dictionaries that we used in our study.

THE TAXONOMY USED IN THE CURRENT STUDY

Based on participants' ratings of self and others, using a large number of person-descriptive terms and a variety of

Table 1. The thirteen languages studied and the 15 dictionaries used

Language	Language family	Dictionary	Region	Human-attribute entries	Total entries	Human-attribute % of total	Ubiquity
Afar	Afro-Asiatic	Parker & Hayward (1985)	North Africa	3,455	9,300	37	81
Hopi	Uto-Aztecan	The Hopi Dictionary Project (1998)	North America	2,725	30,000	9	79
Khoekhoe	Khoisan	Haacke & Eiseb (2002)	Southern Africa	2,592	24,500	11	90
Inuktitut	Eskimo-Aleut	Schneider (1985)	Arctic America	1,816	21,300	9	82
Mara Chin (Lakher)	Sino-Tibetan	Lorrain (1951)	South Asia	1,351	7,000	19	78
Fijian	Austronesian	Capell (1968)	Pacific Islands	1,320	5,800	23	58
Hmong (White)	Hmong-Mien	Heimbach (1979)	Southeast Asia	946	5,100	19	53
Maa	Nilotic	Payne & Ole-Kotikash (2003)	East Africa	779	5,600	14	59
Supyire (Senoufo)	Niger–Congo	Carlson (2003)	West Africa	678	5,500	12	54
Enga	Trans-New-Guinea	Lang (1973)	New Guinea	507	5,200	10	37
Enga	Trans-New-Guinea	Boyd (2016)	New Guinea	308	1,300	24	
Wik-Mungkan	Australian	Kilham et al. (1986)	Australia	382	4,200	9	48
Tsimane	Mosetén–Chon	Gill and Gill (1998)	South America	365	6,200	6	34
Kuna	Chibchan	Holmer (1952)	Central/South America	306	6,100	5	34
Kuna (Guna)	Chibchan	Bodin and Aiban (2012)	Central/South America	228	1,700	13	

Note: The table is adapted from Saucier et al. (2014), with the addition of a Tsimane dictionary and the additional Kuna and Enga dictionaries. Languages are ordered by the number of human-attribute entries, with the multiple Enga and Kuna dictionaries grouped together. Ubiquities (current research) shown for Enga and Kuna are the combined totals across the two dictionaries.

factor-analytic techniques, Goldberg (1990) found the same five-factor structure over the course of three studies. In his first study, Goldberg used 1431 adjective terms from Norman's set of 2800 terms (Norman, 1967) that were classified into 75 categories according to Norman's judgments of their similarities in meaning. In the second study, Goldberg used synonym finders and dictionaries to classify trait adjectives into 133 clusters of quasi-synonyms based on 479 trait adjectives. Further trait adjectives were added, resulting in sets of 566 and 587 terms. In the third study, the set of 479 trait adjectives from the second study was further refined using self-ratings for each of the trait adjectives. Results for each of the 133 synonym clusters from the second study were refined, which produced a set of 339 trait adjectives grouped into 100 synonym clusters. When these 100 clusters were analysed, once again, the five-factor structure emerged: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect/Imagination (coined the Big Five). We employ this taxonomy of clusters from the third study because it has the strongest empirical foundations, building on evidence from the first and the second study. For our research, the primary value of this taxonomy is its empirically derived low-level synonym clusters rather than the high-level Big Five structure.

RESEARCH GOALS AND THE RATIONALE FOR THE CURRENT STUDY

This study has the broad goal of advancing cross-cultural personality research by discovering fundamental concepts of personality that exist regardless of context. Early research into personality structure was predominantly based on relatively homogeneous European and North American societies with languages of common origins. However, the assumptions of previous research on personality structure may not hold when studying personality in more diverse, isolated societies. Therefore, it is necessary to assess the generalizability of personality *concepts* as a precursor to future examination of structure in a diverse range of societies.

While our research is exploratory, we have two specific research goals: (i) detect previously unidentified ubiquitous personality concepts using the cluster–classification approach and (ii) create a set of personality concepts that are ubiquitous across a diverse range of cultures.

Meeting these primary goals will enable more indirect aims to be achieved. For example, discovering the degree of ubiquity and nonubiquity of personality concepts will provide evidence on where humans are located on the continuum between universalism and relativism of personality. The *specific* personality concepts that are present and noticed in widely diverse societies are likely to be deeply embedded in our psychology and so may illuminate core human drives and behaviours. Identifying ubiquitous personality traits may also enable investigation of the across-language aspect of the lexical hypothesis—specifically, that the more important attributes will be present in more languages and the most important ones will be present in nearly all languages.

The current study provides an important alternative approach to the work of Saucier et al. (2014). Their use of single-word terms that are part of a definition minimizes the decisions needed to determine the ubiquity of terms, which increases objectivity. However, this approach may limit the number of ubiquitous concepts that can be identified. Our approach examines personality concepts that are broader than single terms or exactly matched words; we make use of a pre-established and empirically derived taxonomy of English-language individual-difference terms, with quasi-synonyms grouped together into distinct traits. Furthermore, in contrast to Saucier et al., the focus of the current study was on personality concepts rather than the broader set of individual-difference terms.

METHOD

The data from this study, in the form of dictionary translations, categorizations, and scripts used to calculate the correlations are available on the Open Science Framework website: https://osf.io/c6239. Predictions were not preregistered.

Our study includes: (i) a set of 16 857 individual-difference entries from 12 languages, with their indigenous-to-English translations (collated by Saucier et al., 2014); (ii) person-descriptive entries from the dictionary of a 13th language, Tsimane (Gill & Gill, 1998); and (iii) additional dictionaries for both Kuna (Bodin & Aiban, 2012) and Enga languages (Boyd, 2016) (Table 1).

We included an additional dictionary for both the Kuna and the Enga language because of the uncertainty regarding the inclusiveness of the original Kuna and Enga dictionaries. The extra Kuna dictionary contained Kuna entries with Spanish translations. These Spanish translations were then translated into English by two translators, with a third translator checking in detail when the previous two translations did not match closely or were ambiguous.

Initially, the first author (Rater 1) evaluated each of the individual-difference entry translations from the 15 dictionaries for those translations that could be categorized into one or more of the 100 personality synonym clusters. Given the potential for subjectivity in the categorization of some of the definitions, a second rater was then enlisted to categorize two subsets of the definitions. Rater 2 categorized definitions that were considered to be the clearest matches to the meanings of the ubiquitous and near-ubiquitous concepts as initially judged by Rater 1 ($n = 453$). Rater 2 also categorized definitions consisting of a 9% random selection of each language's matched individual-difference terms ($n = 380$).

The translated definitions were evaluated in relation to whether their meaning appeared to be the same as one of the 100 synonym clusters as a whole, even if they did not contain an exact word match. We used the following three criteria to judge whether a definition matched the meaning of a cluster: (i) the core of the definition contained a match to a 100 synonym cluster word-root; (ii) a term at the core of the definition was a close synonym of a word in the cluster; or (iii) the description matched in meaning to the

synonym cluster, despite not having an exact match in words or synonyms of these words. For example, in the case of (i), 'brave, courageous and taking risks for honor' was considered a match to the synonym cluster of Courage (examples of synonym cluster terms are *brave*, *courageous*, and *daring*). Similarly, in the case of (b), *valiant*, with the English-dictionary definition (Yallop, 2006, p. 1316) being 'brave, courageous, or stout-hearted, as persons', was considered a close match to the synonym cluster of Courage, particularly as the definition of *valiant* includes exact terms from the cluster of Courage (i.e. brave and courageous). As another example, *persevering* was considered a match to the synonym cluster of Persistence (e.g. *industrious*, *persistent*, *tenacious*, and *thorough*), with the dictionary definition for *persevere* being 'to persist in anything undertaken' (Yallop, 2006, p. 901). Finally, in the case of (iii), the description 'slow of understanding' was considered a close match to the synonym cluster of Stupidity (e.g. dull, ignorant, and unintelligent). Clearly, of these three criteria, (iii) requires the most subjective judgement.

Many entries were considered too loose a match to be categorized and counted in any tallies. A translated definition was not categorized at all if (i) it was too heterogeneous, (ii) it contained a relevant term that nonetheless was only one part of a broader definition that focused much more on other attributes, or (iii) it was a description of a trait that provokes a reaction in others. For example, for criterion (i), the following English translation of an entry from the Mara Chin language was too diffuse to categorize: 'plucky; daring; brave; tame as a bird; sure; trustworthy; reliable; confident; assured; courageous; insane; foolish; ignorant; mad; silly; idiotic'. The definition contains terms from several different clusters. Other definitions include a term that matches exactly to a synonym cluster term but are not a match in meaning to that synonym cluster. Individual terms within definitions can clearly be different in meaning from the overall definition. For criterion (ii), the example 'to be out of sorts; to be languid; to be slightly unwell' would be too loose a fit to the synonym cluster of Lethargy (e.g. *lethargic* and *sluggish*). Finally, an example for criterion (iii) is the phrase 'cause to insult', indicating that a person does something or has a trait that causes someone else to insult him or her. This phrase suggests a cultural awareness of the person-descriptive concept, but it is not clearly about that particular person.

Definitions that were ambiguously applicable to people, such as *blue* (the colour blue or depressed?) and *cold* (a low temperature or emotionally distant?) were not classified as matches to the 100 clusters. Moreover, the antonyms of a synonym cluster were not counted—for example, *cowardly* was not counted as an indication of Courage.

States and traits were often difficult to distinguish and were both considered to be potential matches to the 100 synonym clusters. For example, *angry* could be a temporary state that most people experience at some time, or it could be used as a trait, said of someone who tends to be angry more often than others. Guided by Angleitner, Ostendorf, and John's (1990) approach to categorizing different types of person descriptors, definitions are considered if they could be used to differentiate one person from another in general

(e.g. short-tempered) or at a single point in time (e.g. angry). In addition, as with Angleitner et al., adjectives (e.g. creative), type nouns (e.g. loser), and attribute nouns (e.g. creativity) were considered.

RESULTS

Matches in meaning

A total of 4237 definitions were matched to synonym clusters by Rater 1 across all 15 dictionaries. Table 2 presents the matches in meaning of indigenous-to-English translations with the 100 synonym clusters as judged by Rater 1. Most translations matched only one synonym cluster—that is, they were a single, or individual, match—indicated by an 'I' in Table 2. However, some translations matched more than one synonym cluster. For these multiple matches, the best or primary match was chosen, also indicated with an I, with the other (secondary) matches indicated by an 'M' (see Appendix A for more detail). Unless specifically noted, ubiquity figures given throughout this article are based on single or primary matches rather than multiple (secondary) matches.

Using Rater 1's categorizations of the definitions from each language, definitions were found to match one of the 100 clusters in 787 out of the 1300 possibilities (13 languages × 100 synonym clusters) or in 61% of the cases, although there is substantial variation across languages (Table 2). Khoekhoe had the most matches (90 out of the 100 synonym clusters), whereas Kuna and Tsimane had the least number of matches (each 34 out of 100), and Enga had the third least (37 out of 100). Due to their low prevalence of matches, Kuna, Tsimane, and Enga exert a strong influence on the overall relative ubiquity. Indeed, there was a substantial relationship between the number of human-attribute terms in a language's dictionary and the number of matches in that language to one of the 100 clusters. The correlation between the number of attribute terms (added in languages with two dictionaries) and the number of matches (0 to 100) was $r(11) = .84$, $p < .001$, 95% CI [0.53, 0.95].

Table 3 shows the ubiquitous (13/13) and near-ubiquitous (12/13) matches for each rater and where the raters were consistent with each other. There clearly is considerable overlap between raters for these concepts, with nine ubiquitous and six additional near-ubiquitous concepts found by both raters.

However, for the set of 9% randomly selected definitions, the overlap between the two raters was smaller. Interrater reliability was calculated using simple joint probability of agreement, given the low probability of choosing the same cluster by chance from 100 clusters. Rater 2's primary match to a synonym cluster was the same as Rater 1's primary match in 242 out of 380, or in 64%, of the cases. If instances are included where a secondary match by Rater 2 matched Rater 1's primary selection, or vice versa, the number of matches between raters increases to 265 out of 380, or to 70%. There were also a further 12% of the cases where what was considered a close primary match by Rater 1 was considered a loose primary match by Rater 2. The lower overlap for the randomly selected

Table 2. Match of synonym clusters to English-translated definitions from the 13 languages

Synonym cluster	Terms	Khoekhoe	Inuktitut	Afar	Mara	Hopi	Fijian	Maa	Supyire	Hmong	Wik-Mungkan	Tsimane	Enga	Kuna	Multiple not included as match	Missing language
Factor 1—Extraversion																
1. Spirit	Spirited, enthusiastic, vivacious, zestful	I	I	I	I	I	M	I	M	I	M	I		M		
2. Gregariousness	Gregarious, extroverted, sociable	I	I	I	I	I	I	I	I	M	I	I	I	I		
3. Playfulness	Playful, adventurous, mischievous, rambunctious	I	I	I	I	I	I	I	I	I	I	I	I	I		
4. Expressiveness	Expressive, communicative, verbal	I	I	I	I	I	I	I	I	M	M	I	I	I		
5. Spontaneity	Spontaneous, carefree, happy-go-lucky	I	I	I	M	I	I	I	M	I	M	M	I	I		
6. Unrestraint	Unrestrained, impetuous, uninhibited	I	I	I	I	I	I	I	I	I	M	M		I		Enga
7. Energy Level	Energetic, active, vigorous	I	I	I	I	I	I	I	I	I	I	I	I	I		
8. Talkativeness	Talkative, verbose, wordy	I	M	I	I	I	I	I	M	I	I	M	I	I		
9. Assertion	Assertive, dominant, forceful	I	I	I	I	I	M	I	I	I	I	M	I	I		
10. Animation	Demonstrative, exhibitionistic, flamboyant	M	I	I	I	I	I	M	I	I	I	M	I	I		
11. Courage	Courageous, brave, daring	I	I	I	I	I	I	I	M	I	I	I	I	I		Supyire
12. Self-esteem	Assured, confident, proud	I	I	I	I	I	I	I	M	I	I	I	I	I		
13. Candour	Direct, frank, straightforward	I	I	I	I	I	M	I	I	I	I	I	I	I		
14. Humour	Humorous, witty	I	I	I	I	I	I	I	I	I	I	I	I	I		
15. Ambition	Ambitious, enterprising, opportunistic	I	I	I	I	I	I	I	I	I	I	I	I	I		
16. Positive Affect	Optimistic, cheerful, jovial, merry	I	I	I	I	I	I	I	I	I	I	I	I	I		
17. Aloofness	Seclusive, unsociable, withdrawn	I	I	I	I	I	I	I	I	I	M	I	I	I		
18. Silence	Silent, quiet, untalkative	I	I	I	I	M	I	I	I	I	I	I	I	I		
19. Reserve	Reserved, detached, secretive	I	I	I	I	I	I	I	I	I	I	I	I	I		
20. Shyness	Shy, bashful, timid	I	I	I	I	I	M	I	I	I	I	I	I			
21. Inhibition	Inhibited, restrained	I	I	I	I	I	I	I	I	I	I	I	I			
22. Unaggressiveness	Unaggressive, unadventurous, uncompetitive	I	I	I	I	I	I	I	I	I	I	I	I			
23. Passivity	Passive, docile, submissive	I	I	I	I	I	I	I	I	I	I	I	I	M		
24. Lethargy	Lethargic, sluggish	I	I	I	I	I	I	I	I	I	I	I	I	I		
25. Pessimism	Pessimistic, bitter, joyless, melancholic, moody, morose, sombre	I	I	I	I	I	I	I	M	I	I	I	I	I		Supyire
Factor 2—Agreeableness																
26. Cooperation	Cooperative, accommodating, agreeable, helpful, patient, peaceful, and reasonable	I	I	I	I	I	M	I	I	I	I	I	I	I		
27. Amiability	Amiable, cordial, friendly, genial, pleasant	I	I	I	I	I	I	I	I	I	I	I	I	I		Wik-Mungkan
28. Empathy	Considerate, kind, sympathetic, trustful, understanding	I	I	I	I	I	I	I	I	M	I	I	I	I		
29. Leniency	Lenient, uncritical, undemanding	M	I	I	I	M	M	I	I	I	I	I	I			
30. Courtesy	Courteous, diplomatic, polite, respectful, tactful	I	I	I	M	I	I	I	I	I	I	I	I			
31. Generosity	Generous, benevolent, charitable	I	I	I	I	I	I	I	I	I	I	M	I	I		

(Continues)

Table 2. (Continued)

Synonym cluster	Terms	Khoekhoe	Inuktitut	Afar	Mara	Hopi	Fijian	Maa	Supyire	Hmong	Wik-Mungkan	Tsimane	Enga	Kuna	Missing language (Multiple not included as match)
32. Flexibility	Flexible, adaptable, obliging	I	I	I	I	I	M	I	I	I	M	I	I	I	
33. Modesty	Modest, humble, selfless, unassuming	I	I	I	I	I	I	M	I	I	I	I	I	I	
34. Morality	Moral, ethical, honest, principled, sincere, truthful	I	I	I	I	I	I	I	I	I	I	I	I		
35. Warmth	Warm, affectionate, compassionate, sentimental	I	I	I	I	I	I	I	I	I		I	M	I	
36. Earthiness	Down-to-earth, folksy, homespun, simple	I	I												
37. Naturalness	Natural, casual, easygoing, informal, relaxed	I	I	I	I	I	I	I	I	M	I	I		I	
38. Belligerence	Antagonistic, argumentative, combative, quarrelsome	I	I	I	I	I	I	I	I	I	I	M	I	I	
39. Overcriticalness	Faultfinding, harsh, unforgiving, unsympathetic	I	I	I	I	M	I	I	I	I	I	I	I	I	
40. Bossiness	Bossy, demanding, domineering, manipulative	I	I	I	I	I	I		I	M	I	I	I	I	
41. Rudeness	Rude, abusive, disrespectful, impolite, impudent, scornful	I	I	I	I	I	I	I	I	I	I	I	I	I	
42. Cruelty	Cruel, ruthless, vindictive	I	I	I	I	I	I	M	I	I	M	I	I	I	
43. Pomposity	Pompous, condescending, smug, snobbish	I	I	I	I	I	I	I	I	M	I	M	M	I	
44. Irritability	Irritable, crabby, cranky, grumpy	I	I	I	I	I	M	I	I	I	I	I	I	I	
45. Conceit	Conceited, boastful, egocentric, egotistical, vain	I	I	I	I	I	I	I	I	I	I	I	I	I	Enga
46. Stubbornness	Stubborn, bullheaded, obstinate	I	I	I	I	I	I	I	I	I	I	I	I	I	
47. Distrust	Distrustful, cynical, sceptical, suspicious	I	I	I	I	I	I	I	I	I	I	I	I	I	Kuna
48. Selfishness	Selfish, greedy, self-indulgent	I	I	I	I	I	I	I	I	I	I	I	I	I	
49. Callousness	Cold, impersonal, insensitive	I	I	I	M	I	I	I	I	I	I	I	I	I	
50. Surliness	Surly, caustic, curt, flippant, gruff	I	I	I	I	I	I	M	I	I	I	I	I	I	
51. Cunning	Cunning, crafty, devious, sly	I	I	I	I	I	I	I	I	I	I	I	I	I	Tsimane
52. Prejudice	Prejudiced, bigoted	I	I	I	I	I	I	I	I	I	I			I	
53. Unfriendliness	Unfriendly, ungracious, unkind	I	I	I	I	I	I	I	I	I	I	I	I	I	
54. Volatility	Volatile, explosive, tempestuous	I	I	I	I	I	I	I	I	I	I	I	I	I	
55. Stinginess	Stingy, miserly	I	I	I	I	I	I	I	I	I	I	I	I	I	
56. Deceit	Deceitful, dishonest, underhanded, unscrupulous	I	I	I	I	I	I	I	I	I	I	I	I	I	
57. Thoughtlessness	Thoughtless, inconsiderate, tactless	I	I	I	I	M	M	I	I	I		I		I	
Factor 3—Conscientiousness															
58. Organization	Organized, orderly, systematic	I	M	I	I	I	I	I	I	I	I			I	
59. Efficiency	Efficient, concise, exacting, fastidious, self-disciplined	I	I	I	I	I	I	I	M	I	I	M		I	

(Continues)

Table 2. (Continued)

Synonym cluster	Terms	Khoekhoe	Inuktitut	Afar	Mara	Hopi	Fijian	Maa	Supyire	Hmong	Wik-Mungkan	Tsimane	Enga	Kuna	Multiple not included as match	Missing language
60. Dependability	Dependable, reliable, responsible	I	I	I	I	I	I	I	I		I					
61. Precision	Precise, meticulous, perfectionistic	I	I	I	I	I	I	I	I	I				I		
62. Persistence	Persistent, industrious, tenacious, thorough	I	I	I	I	I	I	I	I	I			I			
63. Caution	Cautious, careful	I	I	I	I	I	M	M	M	I		I	I	I		
64. Punctuality	Punctual, prompt	I	M	I	I	I	I	I	I	I			M	I		
65. Decisiveness	Decisive, deliberate, firm, purposeful	I	I	I	I	I	I	I	M	I			M			
66. Dignity	Dignified, formal, mannerly	I		I	I	I			I							
67. Predictability	Predictable, consistent, steady	I	I	I	I	I	I	I	I	I						
68. Thrift	Thrifty, economical	I		I	I	I	I									
69. Conventionality	Conventional, traditional	I		I	I	I	I									
70. Logic	Logical, analytical	I	I	I	I	I		I	I			M				
71. Disorganization	Disorganized, haphazard, inefficient, scatterbrained, sloppy, unsystematic	I	I	I	I	I	M	I	I	I						
72. Negligence	Negligent, careless, undependable, unconscientious, unreliable	I	I	I	I	I	I	I	I	I	I			M		
73. Inconsistency	Inconsistent, erratic, unpredictable	I	I	I	I	I	I	I	I	I	I					
74. Forgetfulness	Forgetful, absent-minded	I	I	I	I	I	I	I	I	I	I		I			
75. Recklessness	Reckless, foolhardy, rash	M	I	M	I	I	I	I	I	I	I					
76. Aimlessness	Aimless, unambitious	I	I	I	I	I		I	M	M						
77. Sloth	Slothful, lazy	I	I	I	I	I	I	I	I	I	I	I	I	I		
78. Indecisiveness	Indecisive, wishy-washy	I	I	I	I	I	I	M	I	I	M		I	I		Tsimane
79. Frivolity	Frivolous, extravagant, impractical	M	M	I	I	I	M	I	M	M	M					
80. Nonconformity	Nonconforming, rebellious, unconventional	I	I	I	I	I	I	I	I	I	I					
Factor 4—Emotional Stability																
81. Placidity	Passionless, unexcitable, unemotional	I	I	I	I	I	I	I	I	I	I	I	I	I		
82. Independence	Independent, autonomous, individualistic	I	I	I	I	I	I	M	I	I	M			M		
83. Insecurity	Insecure, defensive, fretful, negativistic, self-critical, self-pitying	I	I	M	I	I	I	I	I	M	I					
84. Fear	Fearful, anxious, nervous	I	I	I	I	I	I	I	I	I	I		I			
85. Instability	Unstable, temperamental, touchy	I	I	I	I	I	I	I	I	I	I		I			
86. Emotionality	Emotional, excitable	I	I	M	I	I	I	I	I	I	M		I			
87. Envy	Envious, jealous	I	I	I	I	I	I	I	M	M	I		I			Hmong
88. Gullibility	Gullible, naïve, suggestible	I	I	I	I	I	M	I	I	I	I	I	I			
89. Intrusiveness	Intrusive, meddlesome, nosey	I	I	I	I	I	I	I	I	I	I	I	I			
Factor 5—Intellect																
90. Intellectuality	Intellectual, contemplative, introspective, meditative, philosophical	I	I	I	M	M	I	M	I	I	I	I	I	I		
91. Depth	Deep and complex	M	I	I	I	M	I	I	I	M	I			I		

(Continues)

Table 2. (Continued)

Synonym cluster	Terms	Khoekhoe	Inuktitut	Afar	Mara	Hopi	Fijian	Maa	Supyire	Hmong	Wik-Mungkan	Tsimane	Enga	Kuna	Missing language / Multiple not included as match
92. Insight	Insightful, foresighted, perceptive	I	I	I	I	I	I	I	I	I		M	I	I	
93. Intelligence	Intelligent, bright, smart	I	I	I	I	I	I	I	I	I	I	I	I	I	
94. Creativity	Creative, artistic, imaginative, innovative, inventive	I	I	M	I	I	I		I						
95. Curiosity	Curious, inquisitive	I	I	I	I	I	I		I						
96. Sophistication	Sophisticated, cosmopolitan, cultured, refined, worldly	I	I	I	I	I		M			M				
97. Shallowness	Shallow, unintellectual, unreflective	M	I	I	M					M				I	
98. Unimaginativeness	Unimaginative, uncreative		M	M											
99. Imperceptiveness	Imperceptive, unobservant	I		I	I	I			I	I	I	I	I	I	
100. Stupidity	Dull, ignorant, unintelligent	I	I	I	I	I	I	I	I	I	I			I	
100	339														
Meaning match to single cluster: Totals		90	82	81	78	79	58	59	54	53	48	34	37	34	787
Meaning match to multiple clusters: Totals		6	5	5	5	5	13	9	11	12	10	9	3	4	97
Single plus multiple		96	87	86	83	84	71	68	65	65	58	43	40	38	884
		Khoekhoe	Inuktitut	Afar	Mara	Hopi	Fijian	Maa	Supyire	Hmong	Wik-Mungkan	Tsimane	Enga	Kuna	

Single   Across the 13 languages: 787/1300 = 60.5%
Single plus multiple   Across the 13 languages: 884/1300 = 68.0%

Note: I = definition matches in meaning to a single synonym cluster or is the primary match (where there was a multiple match and this cluster was chosen as the closest match) in a multiple match (other clusters were then counted as secondary aspects of the multiple match, M).
M = definition matches in meaning to multiple synonym clusters. Note that this cluster is considered a less close match (secondary match) than the cluster that was counted as the primary match (I). Positive Affect was originally labelled more narrowly as Optimism in Goldberg (1990).

Table 3. Ubiquitous and near-ubiquitous matches of synonym clusters to the 13 languages by the two raters

	Rater 1: First author	Rater 2
Ubiquitous (13/13)	Positive Affect	Positive Affect
	Lethargy	Lethargy
	Rudeness	Rudeness
	Stinginess	Stinginess
	Deceit	Deceit
	Sloth	Sloth
	Fear	Fear
	Intelligence	Intelligence
	Stupidity	Stupidity
		Courage (12)
		Pessimism (12)
		Amiability (12)
	Volatility (5)	Irritability (11)
Near ubiquitous (12/13)	Courage (13)	
	Pessimism (13)	
	Amiability (13)	
	Conceit	Conceit
	Cunning	Cunning
	Indecisiveness	Indecisiveness
	Energy Level (11)	
	Stubbornness (11)	
	Envy (11)	

Note: Parentheses indicate the number of languages the concept was found in by the other rater in the cases when the two raters had different results.

subset may partly reflect its greater proportion of ambiguously or poorly worded definitions.

The most ubiquitous matches in meaning

According to Rater 1's matches, there were 10 ubiquitous personality concepts across all 13 languages: Positive Affect (e.g. *cheerful*, *jovial*, *merry*, and *optimistic*—originally labelled more narrowly as Optimism in Goldberg, 1990), Lethargy (e.g. *lethargic* and *sluggish*), Rudeness (e.g. *abusive*, *disrespectful*, *impolite*, *impudent*, *rude*, and *scornful*), Volatility (e.g. *explosive*, *tempestuous*, and *volatile*), Stinginess (e.g. *miserly* and *stingy*), Deceit (e.g. *deceitful, dishonest*, *underhanded*, and *unscrupulous*), Sloth (e.g. *lazy* and *slothful*), Fear (e.g. *anxious*, *fearful*, and *nervous*), Intelligence (e.g. *bright*, *intelligent*, and *smart*), and Stupidity (e.g. *dull*, *ignorant*, and *unintelligent*) (Table 2). However, if both raters are jointly considered, there are nine ubiquitous personality concepts, with Volatility from Rater 1 not found to be ubiquitous by Rater 2, and Irritability, Courage, Pessimism, and Amiability from Rater 2 not found to be ubiquitous by Rater 1 (Table 3).

Using a less stringent threshold for ubiquity—the presence in 12 of the 13 languages—Rater 1 found nine additional personality concepts that qualified: Energy Level (e.g. *active*, *energetic*, and *vigorous*), Courage (e.g. *brave*, *courageous*, and *daring*), Pessimism (e.g. *bitter*, *joyless*, *melancholic*, *moody*, *morose*, *pessimistic*, and *sombre*), Amiability (e.g. *amiable*, *cordial*, *friendly*, *genial*, and

pleasant), Conceit (e.g. *boastful*, *conceited*, *egocentric*, *egotistical*, and *vain*), Stubbornness (e.g. *bullheaded*, *obstinate*, and *stubborn*), Cunning (e.g. *crafty*, *cunning*, *devious*, and *sly*), Indecisiveness (e.g. *indecisive* and *wishy-washy*), and Envy (*envious* and jealous) (Table 4). Considering both raters jointly, there are six additional concepts present in at least 12 out of the 13 languages: Conceit, Cunning, and Indecisiveness (12 out of 13 for each rater), plus Courage, Pessimism, and Amiability (13 out of 13 for Rater 2 and 12 out of 13 for Rater 1) (Table 3).

Ubiquity and the Big Five

The ubiquitous concepts, as judged by the two raters, were from all of the Big Five factors: Agreeableness (three, all at the negative pole), Extraversion (two), Intellect (two), Conscientiousness (one), and Emotional Stability (one). However, the clusters with the highest loadings on each factor (other-ratings in Table 4 of Goldberg, 1990) were not necessarily the most ubiquitous in the current study. In fact, the correlation between relative ubiquity (0–13), as categorized by Rater 1, and factor loading (absolute value) from Goldberg was not significant, $r(98) = .01$, $p = .913$, 95% CI [-0.19, 0.21].

The use of multiple matches

Allowing multiple (secondary) matches increases the overall percentage of the matches by over 7% to 68% (Table 2, last row). Table 4 lists the personality concepts found by Rater 1 to be ubiquitous in the 13 languages as well as those that were near ubiquitous (missing in one language). Although allowing multiple matches increases the overall ubiquity only by ~7%, it increases the number of ubiquitous (Courage, Pessimism, and Envy) or near ubiquitous (e.g. Silence, Empathy, and Belligerence) concepts found by Rater 1 from 19 to 31 (Table 4).

Additional Kuna and Enga dictionaries

The second Kuna and Enga dictionaries contained 17 and 10 matches, respectively, that were not in the original dictionaries, thereby increasing the total number of ubiquitous concepts. At least seven terms found by Saucier et al. (2014) to be present in 11 of the 12 languages (missing in the original Kuna dictionary), were present in the additional Kuna dictionary studied for this research: *dead*, *rich*, *thirsty*, *pain*, *evil*, *fat*, and *gossip*, with the definitions for *short* and *tall* being ambiguous regarding whether they referred to people or only to inanimate objects. Four terms found to be present in 11 of the 12 languages (missing in the original Enga dictionary), were present in the additional Enga dictionary: *happy*, *sleepy*, *love*, and *slow*, with the definition for *quick* being ambiguous regarding whether it referred to people. The inclusion of these extra dictionaries adds to the 28 ubiquitous terms discovered by Saucier et al.

Table 4. Ubiquitous and near-ubiquitous matches of synonym clusters to the 13 languages: Rater 1

	Ubiquitous (13/13)	Near-ubiquitous (12/13) (additional to ubiquitous)
13 Languages Single or primary matches only	Positive Affect, Lethargy, Rudeness, Volatility, Stinginess, Deceit, Sloth, Fear, Intelligence, Stupidity	Energy Level, Courage[a], Pessimism, Amiability, Conceit, Stubbornness, Cunning, Indecisiveness, Envy
13 Languages Multiple (secondary) matches allowed Tally includes single or primary matches	**Courage**, Positive Affect, Lethargy, **Pessimism**, Rudeness, Volatility, Stinginess, Deceit, Sloth, Fear, **Envy**, Intelligence Stupidity	**Spirit**, Energy Level, **Assertion**, **Silence**, **Passivity**, Amiability, **Empathy**, **Warmth**, **Belligerence**, **Overcriticalness**, **Irritability** Conceit, Stubbornness, Cunning, **Caution**, Indecisiveness, **Placidity**, **Insight**

Note: Additional concepts due to the inclusion of multiple (secondary) matches are in bold.[a]The terms *courage* and *bravery* were present in the translations of the terms in the one missing language, Supyire, even though the meaning was not judged to be a primary match. However, it was judged to be a secondary (part of a multiple) match.

DISCUSSION

This study achieved its two specific goals of discovering previously unidentified ubiquitous personality concepts and creating a set of personality concepts that are ubiquitous across a diverse range of cultures. In doing so, it also provided support for the potential utility of the cluster–classification approach.

Key aspects of the lexical hypothesis are that the important human attributes of a society eventually become encoded in its language as single words (Saucier & Goldberg, 1996), and that the most important attributes will be found in the languages of most cultures. What is deemed important will partly be influenced by the social and environmental context, but it is noteworthy that modern-world languages have many single-word concepts that translate accurately into single-word concepts from other modern-world languages. Our findings suggest that this is also the case in small-scale, relatively isolated societies.

In contrast to the first study of this kind (Saucier et al., 2014), which focused on single terms from the English-translated dictionary definitions of individual differences, we focused on personality-trait concepts as indexed by clusters of quasi-synonyms. Our method of matching the meaning of synonym clusters to English definitions yielded nine ubiquitous personality concepts from the 13 languages, found by both raters. Given our reliance on dictionaries to explore indigenous lexicons and the strong possibility that many of our dictionaries were not truly comprehensive, nine may be an underestimate of the actual number of ubiquitous concepts in the languages of the world. Indeed, looser criteria that allow multiple matches and near ubiquity (missing in one language) yielded as many as 31 personality concepts (out of the 100 synonym clusters) for Rater 1. Although some dictionaries may not have been completely inclusive, the second author (who has worked continuously with Tsimane since 1999) specifically checked the Tsimane language for 24 concepts that were found to be missing in Tsimane but were fairly ubiquitous elsewhere. This check, conducted in June 2018 with native Tsimane speakers in Bolivia, confirmed that there

were no single Tsimane words to describe these 24 concepts. In fact, 12 of the 24 concepts had no translations whatsoever that properly captured their essence. This follow-up validation confirms that the low ubiquity of personality concepts in Tsimane is not due to an incomplete Tsimane dictionary but instead accurately reflects the language.

Despite finding ubiquitous personality concepts from each of the Big Five factors, we urge caution when trying to draw conclusions about personality *structure*, Big Five or otherwise, from data on personality concept ubiquity. Rather than the presence or absence of specific ubiquitous concepts from each of the Big Five factors, it is the concept covariation and factor loadings that are required to identify underlying factors. As our results show, there is no correlation between a concept's ubiquity across a broad range of cultures and its loading on its factor in Table 4 of Goldberg (1990). For example, the most ubiquitous concepts within the Extraversion domain (Table 2) were Positive Affect, Energy level, and Courage versus Lethargy and Pessimism. In contrast, the clusters with the highest factor loadings reported by Goldberg were Gregariousness and Spirit versus Silence and Aloofness. If level of ubiquity across many cultures does not correlate with, and is not a clear indication of, a concept's centrality in the factors of one culture's personality structure (in this case, a North American Big Five; Goldberg, 1990), ubiquity may also not be a clear indication of a concept's centrality in the factors of other (non–North American) personality structures. We are wary of concluding that ubiquity, in a set of societies, of *markers* from a particular factor structure will indicate ubiquity of that particular factor structure in the same set of societies.

Our research found only three ubiquitous concepts that overlapped with the ubiquitous terms found by Saucier et al. (2014). These were Lethargy, Fear, and Stupidity. Positive Affect would have been ubiquitous in Saucier et al. (as the term *happy*) if they had the benefit of our additional dictionaries. However, Rudeness, Stinginess, Deceit, Sloth, and Intelligence were ubiquitous in the current study but not in Saucier et al. Courage, Pessimism, Amiability, Conceit, Cunning, and Indecisiveness were near ubiquitous in the

current study (in 12 of 13 languages) but present in less than 11 of the 12 languages in Saucier et al.

Conversely, of the 13 terms found to be ubiquitous by Saucier et al. (2014) that could reasonably be considered personality terms, only three (*afraid*, *tired*, and *stupid*) were found to be ubiquitous in the current study, with *angry* also being ubiquitously matched by each rater in the current study, but to different synonym clusters (Volatility for Rater 1 and Irritability for Rater 2). However, terms ubiquitous in Saucier et al. but not in our study are evaluative terms (*bad*, *good*, and *useless*), emotional states (*ashamed*, *jealous*, and *surprised*), combinations of personality and physical attributes (*strong* and *weak*), and *disobedient*. In addition, the term *gossip* would have been ubiquitous in Saucier et al. if they had had the benefit of our extra dictionaries but the concept was not ubiquitous in our research.

An obvious question is why, if a specific personality term was ubiquitous in the study by Saucier et al. (2014), would it not also be ubiquitous as a concept in the current study? A number of entries that occurred frequently across the 13 languages could not be categorized within the 100 synonym clusters, and explanations for these are listed in Appendix B. However, by way of example, two terms found to be ubiquitous by Saucier et al.—*ashamed* and *surprised*—are more commonly used to describe emotional states than stable traits. These two concepts were therefore not represented in the 100 synonym clusters from Goldberg (1990), given that this taxonomy emphasizes stable personality traits. On the other hand, an argument can be made for the potential of many emotional states to be used in reference to stable traits—for example, people could be considered to have different levels of generalized *shame* and *surprise*.

There are also general methodological reasons why the numbers of ubiquitous terms/concepts in Saucier et al. and in our research are difficult to compare directly. Here, we employed synonym clusters and matches to meaning, and used extra dictionaries, which would be likely to increase the chance of discovering ubiquitous concepts. However, our approach could also decrease the discovery of ubiquitous concepts (i) when a personality term is not considered core to the translation's meaning, (ii) as a consequence of adding an extra language, (iii) by allowing only one match to a cluster per indigenous term (but see Table 4 for multiple matches), and (iv) by confining matches to those found in a fixed, pre-established taxonomy. Nonetheless, our approach allowed us to discover previously unidentified ubiquitous concepts.

Why are some concepts ubiquitous?

What does it mean that across both isolated subsistence-based and urban industrialized populations, people are verbally distinguished by the degree to which they possess these nine characteristics—Positive Affect, Lethargy, Rudeness, Stinginess, Deceit, Sloth, Fear, Intelligence, and Stupidity?

The picture becomes a little clearer if the ubiquitous and near-ubiquitous concepts are grouped into antonyms and close-antonyms: Intelligence–Stupidity, Courage–Fear, Positive Affect–Pessimism, and Energy Level–Lethargy/Sloth. There are also several concepts at the negative pole of the Agreeableness factor, including Stinginess, Rudeness, Deceit, and Conceit, versus Amiability at the positive pole. The following interpretations are necessarily speculative, but below we briefly consider why some of these traits might be especially important to discern in a person, helping to calibrate others' behaviour toward that person (Srivastava, 2010).

Intelligence–Stupidity aligns quite well to the two poles of the Intellect factor in Saucier (1997), emphasizing variation in ability. Level of intelligence, or cleverness, affects many aspects of life in traditional cultures, in all areas of food production and technology manufacturing, navigating dangerous environmental hazards, and managing health problems. It also affects navigation of the social landscape of kin and community, especially regarding maximizing gains from cooperation and minimizing costs of interpersonal conflict. Effectiveness in handling these and many other aspects of traditional life can have serious consequences for well-being, cultural success, and survival.

Courage and Fear may not be exact antonyms, but are likely to be critical attributes in warfare, hunting, and other dangerous activities. Some traditional cultures have within-group physical contests, with participation indicating courage and affecting one's societal standing (Chagnon, 2012). For example, Turkana pastoralists of East Africa mete out punishments and sanctions based on cowardly behaviour during raids (Mathew & Boyd, 2011). However, fear or cowardice might extend one's survival among both sexes in the face of many environmental and societal dangers. Fear and Courage also match quite well to the two poles of the Emotionality domain of a six-factor solution (Ashton et al., 2004).

Stinginess is often highly vilified in subsistence societies that depend on sharing, cooperation, and civic responsibility, particularly when resources are scarce. Disapproval of stinginess seems to be a human universal (Brown, 2000). There also appears to be some consistency across cultures in the way that economic games, measuring aspects of stinginess, are played. For example, the dictator game (Forsythe, Horowitz, Savin, & Sefton, 1994) involves one player unilaterally deciding what proportion of a set pot of money to give to a second player. Across cultures only approximately 40% of the participants offer no money. However, variation in reliance on intragroup cooperation may also affect these results; despite reasonable consistency, there are some clear differences between countries (Forsythe et al., 1994; Hoffman, McCabe, Shachat, & Smith, 1994). More generally, the large number of ubiquitous concepts from the Agreeableness factor may indicate just how important cooperation and teamwork within groups have been in societies, especially in domains of food production, collective defence, and information sharing.

Turning to the broader implications of these results, it seems reasonable to argue that if a concept that signifies variation of a human attribute is created in a society's language, variation in that attribute likely exists between people in that society. This concept would also need to have

sufficient utility to then be maintained through use (e.g. Goldberg, 1982; Norman, 1967), especially in the absence of written language. The details of these two assumptions could benefit from more empirical testing. However, accepting these assumptions for the moment, if an attribute such as Rudeness is ubiquitous (terms for it exist in all languages studied), then (i) variation in the attribute occurs *in all groups of humans* studied and (ii) a term for variation in the attribute has some utility *in all societal contexts* studied. In contrast, nonubiquitous concepts, such as Pomposity, may be missing in some societies because of a lack of trait variation or insufficient utility, among other potential unknown causes.

Future research

Role of trait variation in ubiquity

The level of importance or utility is one likely explanation for differences in ubiquity. However, limited variation in an attribute is another possible cause of a concept's absence in a language, unlikely as this seems for most attributes (e.g. Kajonius & Giolla, 2017; McCrae, 2004). This might be tested by measuring an attribute's between-person variation in societies *without* a single word for the attribute, as compared to societies *with* a single word for the attribute. Traits without single terms could be measured via self- or peer-rating on multiple-word descriptions of the concepts, or where multiple-word descriptions are not possible, via observation. For all of the nonubiquitous attributes documented in the current study, this potential research could help to clarify whether lack of, or substantially restricted, variation is ever a cause of a trait not existing as a single term in a language. Resolving that question might then clear the way for further testing of the lexical hypothesis, in regard to the link between an attribute's importance and its presence as a single word.

The lexical hypothesis

Our results showing the level of ubiquity of the 100 synonym cluster attributes can help to test the across-language aspect of the lexical hypothesis—namely, that the more important attributes will be present as single words in more languages. One approach is to ask people across societies to judge the social importance of a broad range of their language's human-attribute terms, for example by rating how much a trait's presence in a person would determine whether one entered into different types of personal or work-based relationships with that person (Wood, 2015). Each term's importance could be averaged across cultures, or one could focus on relative importance within a culture. Term importance could then be correlated with term ubiquity across cultures to test the across-language aspect of the lexical hypothesis. Additionally, this approach could include traits that are only able to be expressed in multiple-word definitions, to test whether the more important attributes become encoded in a society's language as single words.

The single trait-term criterion

As outlined above, it should be established whether in smaller-scale societies (i) variation exists for most, if not all, attributes that do not have single words and (ii) attributes without single words are judged to be as important as those with single words. This would determine whether the presence of a concept as a single term in a society's language should be the prime criterion for whether it is included for measurement in psycholexical studies on personality structure. If the languages of some societies do not contain single-term concepts for particular human attributes, then multiple-word items could be considered for inclusion in psycholexical research into structure in these societies, if they have similar variation and importance to that of single-word attributes. It may, however, turn out that nearly all of the more socially important attributes of a particular society are encoded as single words. In that case, researchers of personality structure in societies with small lexicons will be faced with an explicit decision: are they interested in determining (i) the covariation of traits that exist in groups of people or (ii) the covariation of traits that are considered socially important enough in a particular society to warrant single-word descriptors? If a group of people moves from one society to another, then the traits deemed important may change along with the single-word attributes, but the covariation of that group's traits may stay the same.

Use of ubiquitous concepts to examine variation in structure

It seems likely that small-scale societies will not always show the same trait covariation as other societies. For example, Gurven et al. (2013) translated the Spanish version of the 44-item Big Five inventory (Benet-Martínez & John, 1998) into Tsimane and in self- and other-ratings found a two-factor structure of personality ("Prosociality" and "Industriousness") rather than the familiar five-factor structure. Future research could measure the covariation of ratings on the core group of ubiquitous personality concepts found here and of the terms found by Saucier et al. (2014), to compare structural personality models across diverse, isolated societies. This may partly solve the problem that cross-cultural comparisons often use personality taxonomies with trait concepts that do not overlap substantially (Fetvadjiev & Van de Vijver, 2017). However, psycholexical studies generally require many more trait terms for people to self-rate on than the number we found to be ubiquitous. Nevertheless, it would be possible to measure the correlations between ubiquitous concepts in this small set and to evaluate the similarity of these correlations across a range of societies and languages. This would provide strong evidence on how universal these correlations are and, by inference, how universal personality structures based on a large number of items are likely to be.

Other personality-trait taxonomies

Our study used Goldberg's (1990) 100 synonym clusters for matching to the English translations of person-descriptive terms from other languages; however, these synonym clusters were not specifically designed for that purpose. To investigate

ubiquity with cluster methodologies in the future, synonyms must be empirically close in meaning, perhaps as judged by multiple raters, and it would be an improvement to have a more comprehensive taxonomy. This taxonomy of synonyms could cover most of the English lexicon that relates to personality attributes, plus include emotional states/traits, like *surprised* and *angry*, and broad evaluative terms, like *good* and *bad* (see Block, 1995; Norman, 1967; Saucier, 1997, regarding considerations for variable inclusion). This may improve on thesauri, which are not always consistent with each other regarding what counts as a synonym (Wood, 2015). A further extension would be to ask linguists and anthropologists to translate this large set of English-language person descriptors into single-word and multiple-word definitions in as many languages as possible. This pre-established taxonomy would enable cross-cultural matching of individual differences between a large number of languages and societies.

Limitations

There was evidence that some of our indigenous dictionaries insufficiently sampled the language, as demonstrated by the additional matches to synonym clusters once the extra Kuna and Enga dictionaries were added. However, sometimes there are no truly comprehensive dictionaries for small-scale societies. In a broader sample of languages, one could attempt to adjust for indicators of variable dictionary quality, such as time dictionary producers spent living within the culture, the number of people who created the dictionary, the range of sources included in the process, and the number of native speakers involved. Moreover, given that dictionaries do not always include all the personality terms in a culture's language, more comprehensive sets of indigenous terms may also be found using a GloCal (global/local) approach (Daouk-Öyry, Zeinoun, Choueiri, & van de Vijver, 2016). This incorporates personality-term identification from spoken personality descriptors, including from laypeople, as well as nondictionary written materials.

Another limitation pertains to the method used for categorizing the matches for Rater 2 as compared to the original matching by Rater 1. Rater 2 received two sets of definitions to match to synonym clusters, one of which was a set of definitions considered to be clear matches to the meanings of the ubiquitous and near-ubiquitous concepts as initially judged by Rater 1. The subsets were used to limit the number of items, the workload, and the timeframe to a reasonable level for Rater 2. The shortcoming is that Rater 2 may have been primed to choose a synonym cluster that they remembered using in many of the groups they had previously coded, leading to an increase in overlap with the synonym clusters chosen by Rater 1. To limit this problem, we randomized the order of presentation so that items were not grouped by synonym cluster for Rater 2. For example, all translations that might match to Stinginess were scattered throughout the subset rather than grouped together. Nonetheless, if this priming effect was going to occur, then the randomized order was unlikely to have eliminated the effect entirely.

CONCLUSIONS

Two raters found that the personality-trait concepts of Positive Affect, Lethargy, Rudeness, Stinginess, Deceit, Sloth, Fear, Intelligence, and Stupidity were ubiquitous across all 13 languages. Six more concepts were found by both raters to be near ubiquitous in at least 12 languages. The cluster–classification method and ubiquitous concepts found here suggest a set of fundamental concepts. Rather than taking a purely universalist or relativist stance, we suggest that there is likely to be a core of universal personality concepts, surrounded by a broader array of nonuniversal concepts. Our research highlights both of these, with concepts found in few cultures being less likely to be universal and those found in most cultures having the potential to be found to be universal upon further investigation.

Ubiquitous concepts are those that are always present in a language, even when other concepts are not. Once it has been more definitively established which concepts are ubiquitous and which are not, in a greater number of languages, this may tell us whether there is a universal order of emergence of personality concepts and, more broadly, individual differences, as discussed in Goldberg (1981).

REFERENCES

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*, i–171.

Angleitner, A., Ostendorf, F., & John, O. P. (1990). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality*, *4*, 89–118. https://doi.org/10.1002/per.2410040204.

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., Di Blas, L., Boies, K., et al. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, *86*, 356–366. https://doi.org/10.1037/0022-3514.86.2.356.

Benet-Martínez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the big five in Spanish and English. *Journal of*

Personality and Social Psychology, 75, 729–750. https://doi.org/10.1037/0022-3514.75.3.729.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187–215. https://doi.org/10.1037/0033-2909.117.2.187.

Bodin, O., & Aiban, W. (2012). *Diccionario escolar: Gunagaya-Español [School dictionary: Gunagaya-Spanish]*. Fondo Mixto Hispano Panameño de Cooperación.

Boyd, R. A. (2016). *Enga dictionary*. Unpublished manuscript.

Breugelmans, S. M., & Poortinga, Y. H. (2006). Emotion without a word: Shame and guilt among Rarámuri Indians and rural Javanese. *Journal of Personality and Social Psychology*, 91, 1111–1122. https://doi.org/10.1037/0022-3514.91.6.1111.

Brown, D. E. (2000). Human universals and their implications. In N. Roughley (Ed.), *Being humans: Anthropological universality and particularity in transdisciplinary perspectives* (pp. 156–174). Walter de Gruyter.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38, 476–506.

Chagnon, N. A. (2012). *The Yanomamo* (6th ed.). Nelson Education.

Daouk-Öyry, L., Zeinoun, P., Choueiri, L., & van de Vijver, F. J. R. (2016). Integrating global and local perspectives in psycholexical studies: A GloCal approach. *Journal of Research in Personality*, 62, 19–28.

Fetvadjiev, V. H., & van de Vijver, F. J. R. (2017). Cross-cultural research. In V. Zeigler-Hill, & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1–12). Springer.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6, 347–369.

Gill, W., & Gill, R. (1998). *Chimane–English dictionary*. New Tribes Mission.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (pp. 141–165), 2. Sage.

Goldberg, L. R. (1982). From ace to zombie: Some explorations in the language of personality. In C. D. Spielberger, & J. N. Butcher (Eds.), *Advances in personality assessment* (pp. 203–234), 1. Lawrence Erlbaum.

Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, 59, 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216.

Gurven, M., von Rueden, C., Massenkoff, M., Kaplan, H., & Lero Vie, M. (2013). How universal is the big five? Testing the five-factor model of personality variation among forager-farmers in the Bolivian Amazon. *Journal of Personality and Social Psychology*, 104, 354–370. https://doi.org/10.1037/a0030841.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, property rights and anonymity in bargaining games. *Games and Economic Behavior*, 7, 346–380.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality*, 2, 171–203.

Kajonius, P., & Giolla, E. M. (2017). Personality traits across countries: Support for similarities rather than differences. *PLoS ONE*, 12, e0179646. https://doi.org/10.1371/journal.pone.0179646.

Klages, L. (1926). *The science of character* (W. H. Johnston, Trans. 1932). Allen & Unwin http://dispater.atspace.com/.

Leising, D., Scharloth, J., Lohse, O., & Wood, D. (2014). What types of terms do people use when describing an individual's personality? *Psychological Science*, 25, 1787–1794. https://doi.org/10.1177/0956797614541285.

Mathew, S., & Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences*, 108, 11375–11380.

McCrae, R. R. (2004). Human nature and culture: A trait perspective. *Journal of Research in Personality*, 38, 3–14.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66, 574–583. https://doi.org/10.1037/h0040291.

Norman, W. T. (1967). *2800 personality trait descriptors: Normative operating characteristics for a university population*. Department of Psychology: University of Michigan.

Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology*, 73, 1296–1312. https://doi.org/10.1037/00223514.73.6.1296.

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 21–50). Guilford.

Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, 69, 847–879. https://doi.org/10.1111/1467-6494.696167.

Saucier, G., Thalmayer, A., & Bel-Bahar, T. (2014). Human attribute concepts: Relative ubiquity across twelve mutually isolated languages. *Journal of Personality and Social Psychology*, 107, 199–216. https://doi.org/10.1037/a0036492.

Srivastava, S. (2010). The five-factor model describes the structure of social perceptions. *Psychological Inquiry*, 21, 69–75.

Wood, D. (2015). Testing the lexical hypothesis: Are socially important traits more densely reflected in the English lexicon? *Journal of Personality and Social Psychology*, 108, 317–335. https://doi.org/10.1037/a0038343.

Yallop, C. (Ed) (2006). *Macquarie Australian encyclopedic dictionary*. Macquarie Library.

APPENDIX A.
CATEGORIZATION OF SINGLE AND MULTIPLE MATCHES

There are some translations that match with more than one synonym cluster; that is, they have multiple matches. For example, if an indigenous-to-English dictionary term has the definition of 'impetuous/excitable man', this could be considered a match to the synonym terms, and by extension to the meaning, of two different synonym clusters: *impetuous* (from the Unrestraint cluster) and *excitable* (from the Emotionality cluster). A key decision when tallying matches to synonym clusters is whether to include these multiple matches, indicated by an M in Table 2. In one sense, it is justifiable to do this as two personality terms that, according to one culture's understanding, are perceived as part of the same concept (e.g. in an isolated indigenous culture) may appear to another culture (e.g. a Westernized, English-speaking culture) as fairly independent. It is also the case that different concepts can be represented by one word. For example, Breugelmans and Poortinga (2006) found that Rarámuri Indians in Mexico had only one word for the different emotions of shame and guilt, despite differentiating between the aspects of shame and guilt in the same way as do other cultures, that have a separate word for each of these concepts.

However, multiple matches mean that some indigenous terms would, in effect, be counted twice. Because of this, where there was a term matched to more than one of the 100 synonym clusters, a choice was made by the first author

regarding which synonym cluster it was most closely matched to, also indicated by an I in Table 2. This cluster was then counted as the primary match (I), with the other clusters that were not considered quite as close counted as the secondary or multiple matches (M). When a language's dictionary contained several different definitions that matched to a particular cluster, but some of those definitions matched only to that single cluster while other definitions matched to that single cluster *plus* other clusters, this cluster was included in Table 2 as a single match (I).

APPENDIX B.
UNCATEGORIZED DICTIONARY ENTRIES

Personality concepts may exist in a language but not have been closely matched to any of the 100 synonym clusters because: (i) they were deemed to be outside of the core factor structure of the Goldberg (1990) taxonomy, based on the original factor analysis decisions (e.g. religious and lusty); (ii) they were very *general* evaluation terms relating to moral character or competence (e.g. bad, good, useless, skilful, mature, and likable), whereas many of the clusters of the Goldberg (1990) Big Five could be considered specific evaluations (e.g. Cruelty, Stupidity, Generosity, and Insight);

(iii) they were religious moral evaluation terms, perhaps influenced in some cases by authors of missionary-created dictionaries changing non-religious moral evaluation terms into those with religious connotations (e.g. *juchaj*, meaning *bad thing someone does* in the Tsimane language, has been extended to the Christian concept of *sin* in a missionary-created dictionary; the Tsimane term *tsij'can*, which literally translates as *in the fire*, has been used by missionaries to describe *Hell*); (iv) they were temporary, emotional states that do not translate readily to dispositions (e.g. shame, regret, grief, lonely, surprised, bewildered, frustrated, startled, and disgusted); (v) they were terms difficult to define (e.g. common sense, distressed, and troubled); (vi) they were too specific (e.g. the fear of being contradicted); (vii) they were based on a behaviour that only implied a trait (e.g. murderer, crying, fighting, complainer, and stutterer); (viii) they were based on a reaction of people to a trait or behaviour (e.g. cause to insult); (ix) they fell loosely among several synonym clusters (e.g. *alert* might loosely match to clusters of Energy Level, Caution, and Intelligence but closely to none); (x) they represented a physical state or trait (e.g. asleep, sick, injured, tired muscles, beautiful, fat, and thin); or (xi) they represented a physical attribute relating to skill (e.g. clumsy and uncoordinated).