Journal Pre-proofs

Big six personality traits in the Africa Long Life Study

Marie Kura, Katrin Rentzsch, Amber G. Thalmayer

Please cite this article as: Kura, M., Rentzsch, K., Thalmayer, A.G., Big six personality traits in the Africa Long Life Study, *Journal of Research in Personality* (2025), doi: https://doi.org/10.1016/j.jrp.2025.104603

Big Six Personality Traits in the Africa Long Life Study

Marie Kura[1], Katrin Rentzsch[2] and Amber G. Thalmayer[3],

[1] Psychologische Hochschule Berlin

[2] Catholic University of Eichstätt-Ingolstadt

[3] Universität Zürich

Contact Information

Correspondence concerning this article should be addressed to Marie Kura, Psychologische Hochschule Berlin, Am Köllnischen Park 2, 10179 Berlin, Germany. Phone number: +49 157 370 5160. Email: m.kura@phb.de.

Sumaya Laher and Tasneem Hassem of the University of the Witwatersrand, Maria Florence, Tracy-Ann Adonis, and Casey Botha of the University of the Western Cape.

Abstract

We explored Big Six personality traits and their correlates among 18-year-olds from Kenya, Namibia, and South Africa (N = 1,746). Established Big-Five and Six-inventories contain culture-specific phrasing and lack fit and measurement invariance in Africa. Using the original maker terms from diverse lexical studies used to build the HEXACO, we created new, single-term Big Six scales with good fit and measurement invariance across countries, which we used to explore concurrent and predictive associations with mental/physical health, and religiosity. Results for Honesty, Agreeableness, Conscientiousness, and Openness replicated findings from North America and Europe, while Extraversion and Emotionality did not, indicating more cultural-specificity. Our method represents a middle path between imported and culture-specific personality research, a promising approach for cross-cultural research.

Keywords: Big Six, cross-cultural personality psychology, personality-outcome associations, Africa Long Life Study

1 Big Six Personality Traits in the Africa Long Life Study[1]

Personality is important to humans worldwide, as indicated by the presence of personality-related terms in every known language (Dixon, 1982; Saucier, Thalmayer, & Bel-Bahar, 2014). Substantiating this notion, empirical research consistently demonstrated that personality traits predict outcomes in practically all relevant life spheres (Ozer & Benet-Martínez, 2006; Soto, 2019). Yet, these results are mainly based on data from Western[2], university-educated samples, which only represent about 11% of the world's population (Arnett, 2008; Henrich et al., 2010; Thalmayer, Toscanelli et al., 2021). Hence, it remains an open question if the same personality traits are equally important for the same outcomes across cultural contexts.

Cultural psychologists agree that personality has both universal aspects that generalize across diverse cultural contexts and more culture-specific aspects that are unique to certain contexts (Thalmayer et al., 2022). However, personality psychology has long struggled with overgeneralizing Western findings and assuming, often uncritically, that they apply universally (Syed, 2024). (Cross)-cultural evidence suggests that some findings assumed to be universal may, in fact, be culture-specific (e.g., Laajaj, 2019). Yet, research in this area remains scarce, leaving open the question of which trait aspects are broadly universal and which are not. For cross-cultural studies, it would be immensely useful to identify ubiquitous trait aspects, that is, aspects that generalize across a wide range of cultural contexts, with respect to their meaning and their role for outcome variables. Establishing this "common ground" in personality traits would enable more valid cross-cultural comparisons and stronger theory testing in personality psychology (Leising et al., 2022). This study aims to identify trait aspects that generalize across diverse cross-cultural contexts, while ensuring their measurement is both psychometrically sound and efficient.

1.1 The Big Six Traits across Cultures

Six-factor models of personality trait structure were proposed as a more cross-cultural alternative, after diverse lexical studies failed to optimally replicate the Big Five (Lee & Ashton, 2008; Saucier, 2009). This work led to the HEXACO inventory platform, including traits of Honesty/Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience, based on convergent results from lexical studies in seven languages; Dutch, French, German, Hungarian, Italian, Korean, Polish (Ashton et al., 2004). Saucier (2009) built on these results to by comparing to a more globally diverse set of studies, and by assessing for HEXACO aspects that were most generalizable, by identifying adjectives that recurred in at least three of the lexical studies that lead to the development of the HEXACO (Ashton et al., 2004), each of which loaded prominently on one Big Six factor. This led to a set of 78 items (see Appendix B in Saucier, Thalmayer, Bel-Bahar et al., 2014) that according to Saucier (2009) can be interpreted as relatively-ubiquitous Big Six markers and therefore have optimal potential to measure a "common ground" in those personality traits across cultural contexts, albeit still mainly western ones. Saucier (2009) referred to this as the "Narrowband cross-language Big Six."; *narrowband,* because Ashton et al. (2004) used variables with highly restrictive selection criteria for their lexical

---

[1] The hypotheses and analyses for this study were preregistered (https://osf.io/25bzs?view_only=73d1df5e17344a3f9302990038c4d65b).

[2] 'Western' is an inexact geographical metaphor on a round globe. We use it here to refer to economically-advantaged Western European countries and English-speaking, majority-European-heritage countries in North America (Canada and North America) and the Pacific (Australia and New Zealand).

studies, as a contrast to the wider selection also reported by Saucier (2009), and used in the development of the similar Questionnaire Big Six platform (Thalmayer et al., 2011).

In the current study, we use the Big Six to identify trait aspects that might be shared across three Sub-Saharan African countries, building on previous research that indicates that the Big Six are a useful framework for cross-cultural comparisons (see also related work by García et al., 2022; Lee & Ashton, 2008, and Thielmann et al., 2020, demonstrating the cross-cultural applicability of the closely related HEXACO model).

When using the Big Six traits outside the contexts in which they were developed and validated, however, it is important to keep in mind that they are imported (*etic*). That is, the six traits capture individual differences that were found to be salient and relevant in predominantly Western contexts. The Big Six traits might not be as valid in Sub-Saharan Africa, as no data from these contexts were involved in identification of the model, and they are unlikely to capture all locally-relevant traits in the region, such as those identified in indigenous lexical studies (Thalmayer, Job et al., 2021, 2022). Nevertheless, using imported models for cross-cultural personality research can be practically useful and can potentially advance personality theory (Thalmayer et al., 2022), for example, by assessing which traits are more or less transferable across contexts and for differences with respect to how traits are expressed (McCrae, 2000) and how they develop (Bleidorn et al., 2013). The current study thus focuses on the potential generalizability and utility of the Big Six traits and their concurrent and predictive validity with respect to established trait-outcome associations in African contexts. Specifically, we build upon previous evidence (Ashton et al.,2004; Saucier, 2009; Saucier et al., 2014), by developing brief Big Six scales that cover the aspects of the Big Six traits that are most generalizable across cultural contexts.

The generalizability of personality traits and their correlates across contexts has been the subject of debate, which has sometimes pitted the extreme of culture-specific traits against that of universal traits (Allik & Realo, 2017; McCrae & Costa, 1997; Thalmayer et al., 2022). Both of these extremes are likely rare, however (Allik & Realo, 2017; Fontaine, 2011). Instead, we find it useful to consider intermediary levels, as summarized by Thalmayer and colleagues (2022). An *existential universal* can be explained and translated everywhere, though meaning and associations may vary. For example, assertiveness may lead to success in competitive cultural contexts but could cause friction where cooperation is valued over competition. *Functional universals* have similar meaning and associations across contexts, but are not commonly used or mentally accessible everywhere. For example, while local terms for Extraversion were scarce or non-existent in three African languages, when prompted, interviewees gave consistent examples of extraverted individuals (Thalmayer, Job, et al., 2021). Rare *accessibility universals* are equally accessible and used the same way in all contexts, and according to Fontaine (2011), should have measurement equivalence. Potential candidates are the Big Two dimensions that emerge consistently in emic, bottom-up lexical studies in diverse languages (Saucier, Thalmayer, Payne, et al., 2014; Saucier, Thalmayer, & Bel-Bahar, 2014; Thalmayer et al., 2024).

Although previous research has convincingly established personality trait-outcome associations, the majority of this research uses North-American and European samples. An open question is whether the Big Six traits are existential, functional, or accessibility universals, that is, the degree to which the concepts are transferable and their associations with outcome variables generalize across contexts. Given that personality traits rarely meet the criteria of accessibility universals and that this is complex to test (similar familiarity, predictive associations and measurement equivalence; Allik & Realo, 2017; Fontaine, 2011; Thalmayer et al., 2022), we suspect that Big Six traits fall somewhere between existential and functional universals. To evaluate this, we adopt a two-step approach: First, we tested if

there were aspects of each Big Six trait that can be measured in a psychometrically-sound way within and across three African contexts. If confirmed, this would suggest that Big Six traits might be existential universals. However, it is important to note that this method offers only indirect evidence of existential universality, a thorough exploration of which would necessitate a bottom-up approach. Instead, our study concentrates on functional universality, that is, the degree to which correlations between Big Six traits and outcomes are consistent across contexts. Consequently, in our second step, we examine to what extent the Big Six traits are functionally universal by exploring if established Big Six-outcome associations generalize across diverse contexts, with regard to three outcome variables of particular practical relevance: mental health, physical health, and religiosity.

1.1.1 Big Six Associations with Mental Health, Physical Health, and Religiosity

In previous meta-analyses, higher Honest/Humility, Agreeableness, Extraversion, and Openness and lower Emotionality were related to better mental health and lower levels of stress and anxiety (Pletzler et al., 2023; Zettler et al., 2020; see also Strickhouser et al., 2017 for related Big-Five-based evidence).

Compared to mental health, associations between the Big Six and physical health tend to be smaller (Pletzer et al., 2023; Strickhouser et al., 2017), mainly with low Emotionality and high Conscientiousness (Pletzer et al., 2023). Meta-analytic results indicate a protective effect of Big-Five Extraversion against mortality (Roberts et al., 2007), suggesting that high Extraversion might be linked to better physical health. However, recent evidence is mixed. Stephan and colleagues (2020) found that Big Five Extraversion was positively associated with self-rated health across eight cohort studies, whereas two meta-analyses found the relation to be non-significant (Pletzer et al., 2023; Strickhouser et al., 2017).[3]

Given that a majority of individuals worldwide identify with a religion (Pew Research Center, 2018), faith plays a substantial role in shaping people's lives, identities, and values (Tarakeshwar et al., 2003). Religiosity has been suggested as a potential source of resilience for both mental and physical health (Berkessel et al., 2021; Kasen et al., 2012; Ronneberg et al., 2016). Previous research has shown that personality and religiosity are interrelated. Specifically, large scale cross-cultural studies have indicated that Big Five Agreeableness and Conscientiousness as well as HEXACO Honesty/Humility robustly predict religiosity across religions and cultural contexts (Aghababaei et al., 2012, 2016; Ashton & Lee, 2019; Gebauer et al., 2014; Saroglou, 2010).

1.1.2 Generalizability of Personality-Outcome Associations to Sub-Saharan Africa

Although associations between the Big Six and mental and physical health have not been tested in Sub-Saharan Africa, related research using other personality models replicates findings from Western samples. For example, Kinyanjui and Sum (2023) found that higher Big Five Neuroticism and lower Agreeableness predicted drug use in a sample of Kenyan students, replicating associations observed in Western samples (Terracciano et al., 2008). Relatedly, Nel et al. (2017) demonstrated that lower Emotional Stability and higher Conscientiousness measured by the South African Personality Inventory were related to higher well-being among South African emerging adults.

---

[3] In our preregistration, we hypothesized that Extraversion would be positively associated with physical health based on Stephan et al. (2020). New evidence from Pletzer et al. (2023), published after our preregistration, contradicts this hypothesis.

With respect to religiosity, Gebauer et al. (2014) reported that Big Six-religiosity associations replicated in South-Africa and Schmitt and Fuller (2015) found that Big-Five-religiosity associations generalized to a sample consisting of aggregated data from Botswana, the Democratic Republic of the Congo, Ethiopia, Morocco, South Africa, Tanzania, Zimbabwe, however, they did not report country-level findings.

These studies suggest that the associations between Big Six and mental health, physical health and religiosity in Western samples might generalize to Sub-Saharan Africa, however there are theoretical reasons and other empirical findings that counter this. It has been argued that cultural context and personality are inseparably intertwined, for example, through social roles, norms, and values, as well as socioeconomic, institutional, and political factors, all of which shape or determine the opportunities for an individuals' personality expression and development (Bleidorn et al., 2013; Clausen, 1968; Havighurst, 1973; Rentfrow & Jokela, 2019; Vygotsky et al., 1997). Because cultural contexts differ with respect to which behaviors are rewarded and punished (Hofstede, 2011) personality tendencies might be differently associated with life outcomes. For example, a study comparing collectivistic (Japan, Mexico, Ghana) to individualistic samples (United States, Germany) reported that nation-moderated associations between Big Five traits and well-being (Schimmack et al.,2002). With respect to religion, Ashton and Lee (2019) demonstrated that zero-order correlations between Big Six traits and religiosity are more pronounced in more religious, compared to non-religious, countries (see also Ezirim et al., 2021). Big Six-outcome associations might conceivably be weaker in less individualistic cultural contexts, as the emphasis on personality traits as enduring qualities that characterize a unique individual is most pronounced in Western countries (Heine, 2012; Henrich, 2020; Singh, 2021). In Sub-Saharan Africa, behaviour and personality are seen as shaped by social expectations and situational context (Triandis et al., 1988, c.f. Oyserman et al., 2002). In contexts with a stronger focus on the situation, Big Six-outcome associations could be weaker. Providing initial support for this idea, Kitayama and Park (2021) found that the well-established association between Conscientiousness and biological health did not hold in a Japanese sample.

Taken together, existing research does not address whether Big Six-outcome associations, as observed in European and North-American samples, are applicable in Sub-Saharan Africa. Testing the generalizability of these traits and their outcome-associations could enrich personality psychology by helping to determine which findings represent "universal truths" about personality, and which are specific to Western, individualistic contexts. This could help untangle basic human tendencies from society-level factors, which are currently intermingled in current personality trait and life-outcome research. To address such question, however, researchers are first confronted with challenges in measuring personality traits in the Sub-Saharan African context.

1.2 Measuring the Big Six in Sub-Saharan Africa

Comparing Big Six correlates across countries requires cross-culturally comparable scales. The Big Six/HEXACO are typically measured with self-rated inventories using short phrases (HEXACO-PI-R, Ashton & Lee, 2009; HEXACO-PI: Lee & Ashton, 2006; HEXACO-60: Ashton & Lee, 2009; also see Thalmayer & Saucier, 2014 for the closely related Big Six Questionnaire). Like all the commonly-used Big Five inventories, the Big Six/HEXACO scales were originally developed in English, and later translated for use in other contexts. This approach is problematic because the resulting inventories inevitably include culture-specific content (e.g., *I would be quite bored by a visit to an art gallery;* Ashton & Lee, 2009). In addition, complexity in content and grammatical structure pose challenges for translation and cross-context adaptation (Thalmayer, Saucier et al., 2021). Unsurprisingly, such inventories show poor

reliability and lack validity outside of Western countries (Laajaj et al., 2019; McCrae & Costa, 1997; Schmitt et al., 2007; Thalmayer et al., 2021; 2022). For example, Thalmayer, Saucier et al., (2021) report very low internal consistencies (α =.26 - .53) for a translated Big Six inventory in a sample of Khoekhoe speakers in Namibia.

Another major challenge in cross-cultural personality assessment is measurement invariance, or psychometric equivalence across groups (Fischer & Karl, 2019). In other words, do the items of a given personality inventory measure the same construct in the same manner in different groups? Several studies have established measurement invariance of HEXACO and related Big Six inventories across countries (Ion et al., 2017; Lee et al., 2018; García et al., 2022; Thielman et al.; 2020; Thalmayer & Saucier, 2014), but only two of these included samples from Sub-Saharan African. García et al. (2022) tested the HEXACO-60 in a large cross-national study across 18 countries. Although the six-factor solution replicated in multigroup confirmatory analysis, results indicated that the factors in the two African samples (Togo and Senegal) deviated from those in the other countries included in the study. Moreover, internal consistency was lowest in Togo and Senegal.

Another large cross-national study by Thalmayer and Saucier (2014) tested the Big Six Questionnaire across 33 countries, including three African subsamples. Configural invariance was established, however, according to some experts, comparison of correlations should only be made across groups if metric invariance holds (i.e., factor loadings are equivalent across groups; Fischer & Karl, 2019). As we aim to explore and compare correlations of the Big Six across countries, metrically-invariant scales are necessary for the current study.

Taken together, previous evidence indicates that existing HEXACO/Big Six inventories are not well suited for Sub-Saharan Africa. Therefore, we took a novel approach to assess Big Six traits that matches that used in bottom-up lexical studies of personality traits, and that has been proposed for other types of cross-cultural research (e.g., Romaneo et al., 2023). Specifically, we use single-term marker items (e.g., *shy* or *cheerful*) combined with a simple, standardized item stem that makes clear our interest in characteristic, trait-like tendencies (*I tend to* [*be*]…). The marker items administered in the Africa Long Life Study was derived from the lexical studies originally used to develop the Big Six (Ashton et al., 2004; Saucier, 2009; see Cross-Language Six items in Appendix B in Saucier, Thalmayer, Bel-Bahar et al., 2014). Using these items with a standardized stem avoids many problems associated with importing inventories (e.g., Thalmayer et al., 2022) while making use of the strength of the Big Six: its basis in ubiquitous denominators from globally-diverse lexical research (Ashton et al., 2004; Saucier, 2009). A recent study in an Italian sample (Romano et al., 2023) demonstrated that single adjective HEXACO markers have the potential to form scales with excellent psychometric properties. While the marker items used in their study primarily originated from Italian and English HEXACO studies, and were intentionally specific to their context of interest, our item set draws on a maximally-diverse range of data with the intention to provide a basis for equivalently measuring personality traits across global contexts.

1.3 The Present Study

Prior Big Six literature has two major limitations. First, contexts outside the West are underrepresented, particularly in studies assessing correlates, which leaves the generalizability of existing findings inconclusive. Second, the few studies that included samples from Sub-Saharan Africa used imported personality questionnaires with problematic psychometric properties and reasons to suspect misunderstanding of item content.

To overcome these limitations, we seek to draw valid comparisons between existing Big Six literature and African samples using a rigorous and contextualized approach. Our goal was to create brief, measurement-invariant Big Six marker scales with at least acceptable measurement properties for use in the Africa Long Life Study and for drawing comparison between samples in Kenya, Namibia, and South Africa. Our approach was pragmatic, in that we aimed to identify a subset of items from the original Big Six item pool best suited as common denominators across the three countries. Because the scales created in this study will be used in an ongoing longitudinal study with limited space in a larger item battery, our aim was to create short, yet psychometrically sound scales. Overall, our pragmatic approach to scale construction contrasts with a theory-driven approach, which would emphasize construct validity.

Short scales are commonly used in personality psychology and even (ultra-)brief scales with two to five items have been shown to assess personality traits with high psychometric quality and to predict external criteria at comparable or even larger effect size to longer scale versions, even in the face of lower internal consistency (Oshio et al., 2014; Sandy et al., 2014; Rammstedt et al., 2020; Thalmayer et al., 2011; Ziegler et al. 2014).

To explore to what extent the Big Six traits are applicable in the three countries in the sense of existential universals (explainable and translatable), we evaluated the resulting marker scales with respect to model fit and internal consistency. In addition, we used the resulting marker scales to study associations between the Big Six traits and mental health, physical health, and religiosity within and between the countries. To test the Big Six traits for functional universality, we examined if these associations correspond to those found in previous meta-analyses and large-scale studies based on mostly Western samples, which are summarized in Table 1. Note that although we formulated five to nine hypotheses to predict the relationships between each trait and the outcome (sub-)scales, we specified only two hypotheses for Openness. This decision was informed by prior research, which suggests that Openness holds limited criterion validity in relation to the outcome variables selected for this study (see Ashton & Lee, 2019; Zettler et al. 2020).

2 Methods

To answer our research questions, we used data from the first, second, and fifth wave (W1, W2, and W5) of the Africa Long Life Study (ALLS), a longitudinal study on emerging adults in Kenya, Namibia, and South Africa in which co-investigators from each country helped shape the research questions, materials, and recruitment protocols (Thalmayer et al., 2024). Ethical approval was obtained for the ALLS as a whole by the partner universities, specifically, the Catholic University of Eastern Africa, the University of Namibia, the University of the Free State, the University of the Witwatersrand, and the University of the Western Cape, as well as at the national level in Namibia and at the national and county levels in Kenya. In Namibia and South Africa, authorization from regional counselors and educational officers was obtained prior to recruitment.

The hypotheses and analyses were preregistered https://osf.io/25bzs?view_only=73d1df5e17344a3f9302990038c4d65b. All data and research materials, including the codebook, the anonymized dataset, and the R scripts, as well as an overview on all deviations from our preregistration (see Willroth & Atherthon, 2024) are available at https://osf.io/qe5hn/files/osfstorage?view_only=745452ec7f7c4a2eb851e0339c902498.

2.1 Participants

In each of the three countries, 18-year-old participants were recruited in 2022 for a 5-year longitudinal study with two data-collection waves per year. The inclusion criteria were the ability to read English, to read the informed consent letter, to consent to participate in the study, and to provide detailed contact information to be reached six months later.

Because we used Big Six items from both W1 and W2, we only included participants whose surveys at the two time points could be linked (N = 2,158) and who provided personality information (140 participants from Kenya, 234 from Namibia, and 38 from South Africa did not provide any response to the Big Six item at Wave 1) The resulting analytical sample of N = 1,746 (M_{age} = 18.24, SD = 0.82) is highly diverse, including over three dozen ethnolinguistic groups. Sample characteristics for the analytic sample are reported in Table 2. More details are provided in Thalmayer et al. (2024).

2.2 Procedure

The language of the surveys was English, however, translation of instructions and some measures in local languages (not including personality items) were available to the research assistants to support their participants. Surveys were administered electronically on an ipad or on paper and pencil by research assistants.

2.3 Measures

2.3.1 Big Six Marker Scales

The Big Six marker items administered in the ALLS were based on original work to develop the HEXACO model (Ashton et al., 2004), specifically Saucier's (2009) identification of adjectives that recurred in the majority of the six languages investigated (Dutch, French, German, Hungarian, Italian, Korean, and Polish).

In order to select a subset of the 78 English-language items that would be understood by participants in Kenya, Namibia, and South Africa, ALLS co-investigators and research assistants from the three countries reviewed the list and identified terms that were unknown to them and/or likely to be unknown to most 18-year-olds in their countries. Items were excluded if multiple colleagues from more than one country classified them as not understandable. The resulting item set contained 56 marker items (7 to 11 per trait), which are displayed in Table 4. These marker items contained mainly adjectives and one verb (boast). While the original terms derived from the lexical work by Ashton et al. (2004) was boastful, the item was modified here to improve understandability based on feedback from local team members. Fifteen of the marker items were administered at W1 and 41 at W2, which meant a one-to-six-month interval between the responses. At both waves, the instruction was: *How accurate is each statement in describing your typical behavior or attitudes. Do you tend to (be)…?*. The 6-point response scale ranged from 1 (*very inaccurate)* to 6 (*very accurate)*. To accommodate potential language comprehension issues, an alternative response option, "I do not know this word", was available. "I do not know this word" responses are coded as 777 in the dataset, and were treated as NAs in all subsequent analyses.

2.3.2 Correlates

We included measures of mental health, physical health and religiosity as potential correlates of the Big Six traits, all of which were assessed at W1 and W5.

Mental Health. Mental Health was assessed with the International Mental Health Assessment (IMHA, Thalmayer et al., 2023). The IMHA provides hierarchical information on three levels. First, a P-Factor of general psychopathology reflects an individual's level of overall dysfunction and proneness towards disorder (Caspi et al., 2014; Conway, 2019). Second, broader spectra, including an individual's inclination to: experience internal emotions (internalizing tendencies), as reflected by for example anxiety and depression; and to direct their emotions outwardly (externalizing tendencies), in the form of substance (ab)use and anger issues. On the third and lowest level are narrower subscales for common psychological and behavioral health categories, of which we used those, for which there was evidence suggesting an association with the Big Six (see Table 1): Depression, Anxiety, Substance Use/Abuse, Anger, Life Stress, and Interpersonal and Partner Conflicts. These categories were

rated in terms of frequency experienced over the last month (1: *not in the last month or never* to 7: *daily or almost daily*).

We computed mean scores for P-factor, Externalizing and Internalizing Spectra, and all subscales with higher scores representing lower levels of mental health.

 Physical Health. Physical Health was measured with the single item general self-rated health (GSRH; DeSalvo et al., 2006), using the query *In general, would you say your health is...* rated on a 5-point scale from 1 (*excellent)* to 5 (*poor)*. We inverted the responses so that high scores reflect high levels of physical health.

Religiosity. Religiosity was assessed with the five-item Duke University Religion Index (DUREL; Koenig & Büssing, 2010), which was originally designed to measure three dimensions of religiosity and religious involvement: organizational religious activity, non-organizational religious activity, and intrinsic religiosity (or subjective religiosity indicated by for example experiencing a "presence of the Divine").

The results of an extensive cross-cultural study (Toscanelli et al., 2022) suggest that organizational religious activity (attending religious services/meetings) and non-organizational religious activity (spending time meditating or praying) can be combined into a single religious activity factor with the resulting two-factor structure capturing individual differences in religiosity better in many countries including Kenya and Namibia (Toscanelli et al., 2022). The first factor, religious activity, consisted of two items: *How often do you attend church, mosque, temple, or other religious services/meetings?* and *How often do you spend time in private religious activities, such as prayer, meditation, or study of religious scriptures*?. Religious activity items were answered on a 6-point scale (1 = *never*, 2 = *once a year or less*, 3 = a *few times a year*, 4 = a *few times a month*, 5 = o*nce a week*, 6 = *more than once a week*). The intrinsic religiosity items *In my life, I experience the presence of the Divine.*, *My religious beliefs are what really lie behind my whole approach to lif*e., *and I try hard to carry my religion over into all other dealings in life*. were answered on a five-point-scale: 1 = *definitely untrue*; 2 = *somewhat untrue*; 3 = *neither true nor untrue*; 4 = *somewhat true*; 5 = *definitely true*. Based on feedback from the ALLS team, wording was adjusted for applicability: instead of asking participants how often they "attend church", they were asked how often they attend a "church, mosque, or temple" and instead of asking how often they engage in "bible study", they were asked how often they spend time studying "religious scriptures".

We used the sum scores across all corresponding items for both of these subscales for all subsequent analyses.

2.3.3 Control Variables

Home Language. Home language was assessed at W1 with an open answer, the item wording was adjusted to local norms. To maintain anonymity, we grouped languages into broader language family categories (see Thalmayer et al., 2024, for details).

Gender. Gender was assessed though the item *What is your gender?* at W1. Response options were *Female* (coded as 0), *Male* (coded as 1), and *Other* (coded as 2).

Socioeconomic Background. Socioeconomic background was measured through the Family Affluence Scale (FAS II, Boyce et al., 2005) at W1. The FAS II is a commonly used measure of family affluence or socioeconomic status among adolescents. It includes four indicators of material family wealth (unshared bedroom, household car ownership, number of computers in household, number of family holidays). Response options ranged from 0 (*no/none/never*) to 2 (*two/twice*) or 3(*more than two/more than twice*). Responses were summed for an overall score.

2.4 Analyses

Analyses were preregistered on OSF (https://osf.io/25bzs?view_only=73d1df5e17344a3f9302990038c4d65b). Our analytical plan consisted of four major steps: First, we selected psychometrically-sound items from the Big Six item pool (Table 3) in order to create scales that were measurement invariant across the three samples. For this purpose, we used an item-selection algorithm (Schultze, 2017). Second, measurement invariance of the resulting scales was established through multigroup confirmatory factor analyses (CFA). Based on this, we made some adjustments to the item composition of the scales. Third, we evaluated the final scales in each country in terms of model fit and reliability. Fourth, we used the scales to compute correlations between the Big Six traits and mental health, physical health, and religiosity in each country. We then compared the correlational patterns to hypotheses derived from in existing literature, mainly based on Western samples.

2.4.1 Item Selection

Item Selection from the Initial Item Pool due to Redundancy and Missingness. Before applying the item-selection algorithm, we removed items from the initial item pool according to preregistered criteria. Specifically, to avoid redundancy in the resulting scales, we checked if any of the items from the Big Six item pool were intercorrelated at $r \geq |.80|$. In addition, we checked if items should be excluded due to systematic missingness. Participants had the option to respond with *I don't know this word*. This might lead to values not missing at random (NMAR) which in turn can bias the maximum likelihood estimation's results. To prevent this, we checked for items with more than 25% *I don't know this word* responses. Moreover, we aimed to exclude items for which the *I don't know this word* answer was correlated to a meaningful extent with home language and socioeconomic background. However, due to the high number of home languages present in the data (55 total), assessing the association with missingness was not feasible.

Brute Force Item-Selection Algorithm. Within the resulting item pool, we applied an item-selection algorithm in order to select an optimal subset of items per Big Six trait with the potential to be measurement invariant across countries. Using a selection algorithm bears the advantage of simultaneously evaluating psychometric properties and overall model parameters such as model fit. Previous studies successfully applied similar strategies to create cross-cultural measurement invariant personality scales (Jankowsky et al., 2020). We applied the bruteforce algorithm from the stuart R package (Schultze & Schüller, 2022) to test and compare all possible item subsets per scale. We constrained the bruteforce algorithm to select five items per scale and to identify the (close-to) optimal solution in terms of model fit and reliability when metric invariance is assumed. In doing so, we aimed to create brief, yet psychometrically sound scales (Rammstedt et al., 2020; Thalmayer et al., 2011).

To avoid overfitting, we applied Monte Carlo Cross Validation (Picard & Cook, 1984), randomly extracting 50 subsamples (with replacement), each comprising half the size of the total sample. Subsequently, we applied the item-selection algorithm in each of these subsamples. In the final step, we chose the subset of items for each trait that showed the highest frequency of occurrence across the 50 iterations.

2.4.2 Measurement Invariance

We evaluated the item sets with the highest frequency of occurrence across the 50 iterations with respect to measurement invariance in the full sample. This allowed us to make use of the full sample while avoiding overfitting. Measurement invariance of the marker scales was established through multigroup confirmatory factor analysis using the lavaan (Rosseel, 2012) and semTool (Jorgensen et al., 2022) R packages (R version 4.3.1). This was done separately for each trait. To identify models, we applied effects coding to the item loadings. That is, we constrained the factor loadings of the items for each Big Six trait to average one (Card & Little, 2007). We did not reverse score any items before estimating the measurement models, in order to explore the direction of the loadings freely. This is important because the items have never been used in our study population, hence *a priori* hypotheses regarding the direction of loadings were not justified (but see Saucier, 2009, who reports item loading directions for other populations). Missing data was handled using full information maximum likelihood estimation.

To test configural measurement invariance, we first estimated a model where loadings and intercepts were allowed to vary between countries for each trait. In these initial configural models,

some items had incongruent loading patterns across countries; in other words, positive loadings on the trait in one country and negative loadings in another. Because metric invariance requires equal factor loadings across countries, we removed items with incongruent loading patterns. We then compared the change in CFI of three increasingly restrictive measurement models for each trait; configural, metric (factor loadings constrained to equality), and scalar (factor loadings and intercepts constrained to equality). According to pre-registered criteria, we used a cutoff of .01 for the change in CFI between models (Cheung & Rensvold, 2002) as our criterion for retaining a more restrictive model.

2.4.3 Evaluation of the Final Big Six Marker Scales

The final scales were evaluated by assessing scale properties of the configural measurement models in each country. We used the CFI and RMSEA in each country as indicators of model fit. Hopwood and Donnellan (2010) found that many personality models do not meet standard fit benchmarks, even when they are administered in the cultural contexts they were derived from. Therefore, we adopted their relatively liberal cutoffs as more reasonable than benchmarks derived in other domains: CFI > .90 and RMSEA < .10. We computed Mc Donald's ω as an indicator of the scale's reliability and consider an ω ≥ .70 as sufficient (see also Jankowsky et al., 2020; Olaru et al., 2019).

2.4.4 Correlations with Mental Health, Physical Health, and Religiosity

For each of the resulting Big Six marker scales we computed separate mean scores per country. Items with negative loadings in the multigroup CFA were reversed before computing mean scores. Big Six mean scores were used to compute bivariate correlations between the Big Six and the mean scores for mental health, physical health, and religiosity. We then assessed if the correlations in each country corresponded to those hypothesized in Table 1 in terms of direction and significance ($p < .01$). A sensitivity power analysis conducted in G*Power 3.1 indicated that the smallest of the three country-subsamples used in this study ($n = 503$, South Africa) would provide 99.82% power to detect a trait-outcome correlation with *an r* of .23, assuming a two-tailed test and an alpha of .01. For reference, Soto (2019) reported a mean effect size of .23 across 78 previously published trait outcome associations. Hence, our analytical sample should be sufficiently powered to test the hypothesized trait-outcome associations. Furthermore, because previous studies have indicated that country differences in personality correlates could partly be attributable to gender and social position (García et al., 2022; Zettler et al., 2020), we also controlled for gender and family affluence in our analyses. To do so, we regressed each outcome variable on one personality trait, gender, and family affluence as predictors. These additional regressions were computed separately per country.

In addition, to examining bivariate correlations between each Big Six trait and the outcome variables, we examined if the trait-outcome associations indicated by the bivariate correlations remained stable when accounting for all Big Six traits. In doing so, we computed multiple regressions for each outcome variable, with all Big Six traits as predictors.

Finally, to test the predictive validity of our scales, we examined whether associations between the Big Six traits and outcomes measured at Wave 1 and 2 generalized to the outcomes assessed at Wave 5, two years later.

3 Results

Internal consistencies and descriptive statistics for all measures are reported in Table 4; intercorrelations are displayed in the Supplemental Tables A1-A4.

We aimed to establish measurement-invariant Big Six marker scales in Kenya, Namibia, and South Africa to compare Big Six-outcome correlations between the three countries. In the following subsections, we describe the Big Six marker scales that were developed based on the item-selection algorithm, we present the results of measurement invariance testing and the correlations between the Big Six trait scales and the outcome variables.

3.1 Item Selection from the Big Six Item Pool

According to preregistered criteria, we removed items from the initial Big Six item pool to obtain psychometrically-sound scales. We excluded three items because more than 25% of participants indicated not knowing the words (*solitary, mild*, *industrious*). We excluded four items because their not-known responses were correlated with socioeconomic background at $r \geq .|10|$ and $p < .01$ (*vulnerable, introverted, hot-headed,* and *precise).* All remaining items were intercorrelated below .80, therefore, no items had to be removed due to redundancy.

We applied the bruteforce item selection algorithm to the adjusted pool of 49 items, resulting in the scales presented in Table 5, which make use of 30 items. The maximum frequency of runs across which the selected item set for each trait emerged ranged from six to 28 out of a possible total of 50.[4]

3.2 Measurement Invariance of the Big Six Marker Scales

We computed a multigroup CFA for each trait to explore the loading patterns of the items selected through the bruteforce algorithm. This allowed us to identify items with inconsistent loading patterns across countries, which might lead to non-invariance. Based on the results, we removed three items with inconsistent loadings: S*ly* loaded positively on Honesty in Namibia ($\lambda = 0.14$, $p = .506$), but negatively in Kenya ($\lambda = -0.62$, $p = .005$) and South Africa ($\lambda = -0.77$, $p = .042$); *shy* loaded negatively on Extraversion in Kenya ($\lambda = -0.90$, $p = .009$) and South Africa ($\lambda = - 1.45$, $p < .001$), but positively in Namibia ($\lambda = 0.53$, $p = .085$); *irritable* loaded negatively on Agreeableness in Kenya ($\lambda = -1.219$, $p = .001$), and South Africa ($\lambda = -0.25$, $p = .326$), but positively in Namibia ($\lambda = 0.11$, $p = .616$).

We used these adjusted scales including 27 total items for all subsequent analyses. That is, we again computed a multigroup CFA for each trait with the adjusted scales. To test for measurement invariance, we compared three increasingly restrictive measurement models with the multigroup CFA. Based on the configural model for each Big Six trait, we obtained item loadings and indicators of model fit and internal consistency.

Model fit and comparisons of the models with different equality constraints are displayed in Table 6. We established metric invariance for all scales except for Emotionality and Extraversion which

---

[4] For the Openness scale, two item sets with the same number of iterations (six) emerged. We evaluated both sets through multigroup CFA and chose the one which was measurement invariant across countries. The other item set was non-invariant because it included one item (*ironic)* which loaded in different directions across groups.

both only reached configural invariance. Even though the criterium for metric invariance (CFI Δ < .01) was not reached for Extraversion and Emotionality, the metric measurement models for these traits demonstrated satisfactory fit according to our prespecified criteria (RMSEA < .10, CFI > .90).

Factor loadings obtained from the configural measurement models with the adjusted scales for each trait are displayed in Table 7. A surprise in the factor loadings is that *passive* loaded positively on Extraversion in all three countries, in contrast to negative loadings in prior studies (Ashton, 2004; Saucier, 2009). Similarly, while *emotional,* loaded positive on Emotionality in prior studies (Ashton, 2004; Saucier, 2009), it loaded in the opposite direction in our samples. Post-hoc analyses indicated that the unexpected item loadings of *emotional* and *passive* were not caused by specific response patterns (Supplement B). The loading patterns for all other items were congruent with prior evidence (Ashton, 2004; Saucier, 2009).

3.3 Evaluation of the Final Big Six Marker Scales

The final Emotionality scale only includes reverse-keyed indicators. Three out of five Conscientiousness items were reverse-keyed whereas the Honesty and Agreeableness and Openness marker scale only include forward-keyed indicators. Note, that for Openness, the original item pool only included positively-keyed items, but for Honesty, Extraversion, Agreeableness, and Conscientiousness, at least three reverse-keyed indicators were available (see Table 3).

As displayed in Table 8, the configural measurement models for each trait used for the multigroup CFAs with the adjusted scales revealed good to satisfactory fit (CFIs > .954, RMEAs < .077), with the exception of Conscientiousness, which fit poorly in the Namibian subsample (CFI in Namibia = .861). The internal consistencies (see Table 3), however, were relatively low (most ωs < .70 in each of the countries), particularly for Extraversion (Kenya: ω = .51; Namibia: ω= .31; South Africa: ω = .65). In contrast, Openness was the scale with the highest internal consistency (Kenya: ω = .68; Namibia: ω = .69; South Africa: ω = .67).

3.4 Correlations with Mental Health, Physical Health, and Religiosity

After establishing measurement invariance, we used the final items set of each Big Six trait to compute the corresponding marker scale scores to further examine associations between the Big Six traits and the outcome measures of mental and physical health and religiosity in every country. Before computing the Big Six mean scale scores, we recoded all Emotionality and Conscientiousness items and items with negative loadings so that higher mean scores reflect higher levels of the trait. The resulting bivariate correlations are displayed in Table 9, exact *p*-values and confidence intervals are displayed in Supplemental Table A5. Note that correlates of Emotionality and Extraversion should be interpreted separately for each country because the scales lack metric invariance which, according to common

Overall, the personality-outcome correlations were of small to medium magnitude (IrI range = .00 - .23, IrI average = .08). The correlational patterns were congruent across countries in that we found no correlations to be significantly positive in one country and significantly negative in another. Twenty-seven of the 42 hypothesized correlations were observed in at least one sample. Most of the observed associations in Kenya, Namibia, and South Africa were in line with theory and previous findings from mostly Western samples, although some did not reach the pre-registered significance level of *p* < .01 in one or more of our country samples. A schematic display of hypothesized and observed correlations can be found in Table 1.

In Kenya and Namibia, Honesty was associated with higher levels of mental health (IMHA), that is, Honesty correlated negatively with the IMHA p-factor (Kenya: r = -,18 p < .001; Namibia: r = -.11, p = .005) and the Interpersonal Conflict subscale (Kenya: r = -.16 p < .001; Namibia: r = -.07, p = .063). Also in line with our hypothesis, we found a significant positive association between Honesty and religious activity in Kenya (Kenya: r = .23, p < .001; Namibia: r = .04, p = .329, South Africa: r = .07, p = .115).

Emotionality had fewer significant associations than expected, among the hypothesized associations, only a negative correlation with Intrinsic Religiosity in Namibia (r = -.13, p = .001) and a negative correlation with physical health in South Africa (r = -.12, p = .008) were significant. Contrary to our hypotheses, Emotionality was not linked to mental or physical health in either of the other countries.

As expected, higher Extraversion levels were related to higher substance abuse (r = -.16, p = .< 001), and lower life stress (r = -.12, p = .008), anxiety (r = -.14, p < .001), and depression (r = -.21, p= .001) in South Africa. The hypothesized positive association between Extraversion and physical health was also confirmed in South Africa (r =. 18, p < 001). But Extraversion-mental and physical health correlations were non-significant and close to zero in Kenya and Namibia.

Corroborating our hypotheses, high Agreeableness was consistently associated with lower scores on anger, substance abuse, and externalizing problems in Kenya and Namibia (rs between -.17 - .10, ps between <.001 and .008) and higher levels of religiosity in all three countries. Notably, Agreeableness was not significantly related to any outcome variables in South Africa.

As expected, high Conscientiousness was related to lower overall mental health problems and externalizing problems (rs btween -.18 and -.14, ps between < .001 and .002) in all three countries. Also in line with our hypotheses, Conscientiousness was positively related to both intrinsic and extrinsic religiosity in Keya and South Africa (rs btween .12-and .15, ps between .001 and .005).

Openness was linked to better mental health across the three countries, however, we did not find the hypothesized negative associations with life stress and anxiety. In Kenya and South Africa, high Openness was significantly associated with higher physical health (Kenya: r = .15, p <.001; South Africa: r = .16, p = .001) and religiosity (rs between .12 and .19, ps < .01), for which we had no prior hypotheses.

As displayed in Supplementary Table A6, the Big Six-outcome associations remain stable after controlling for socioeconomic status and gender. Similar associations were observed for the outcome criteria measured at W5 (see Supplementary Table A8-10). In addition, as shown in Table 1 and Supplementary Table A9, patterns of Big-Six-outcome associations in the multiple regressions were similar to the bivariate correlations. Yet, unsurprisingly, after controlling for the other traits, some regression weights were no longer significant. Notably, Conscientiousness remained a significant predictor for both mental and physical health at Wave 1 and Wave 5 (see Table 1 and Table A10), even after controlling for the other traits.

Post Hoc Analyses

Because the Emotionality and Extraversion scales were problematic (lack of metric invariance, incongruent and contradictory item loadings), we explored if we could create scales with better psychometric properties and measurement invariance for these traits, drawing from the items chosen most frequently in the item sets selected across the 50 runs (see Supplementary Table C1). We did so to address the pragmatic need for usable scales in the ALLS specifically, and for testing in other samples

generally. To do this, we identified the items included in the three most frequently selected item sets across the 50 runs, and explored if we could combine them into better scales for these two traits separately. This approach offers more flexibility compared to our original, algorithm-based item-selection approach, which relies on a prespecified number of items per scale. Detailed results are reported in Supplement C. For Emotionality, the resulting scale fit worse than the original scale selected through the item selection algorithm but was metrically invariant and had satisfactory internal consistency. The new Emotionality scale showed some of the expected negative associations with mental and physical health, however, some of these associations were not significant across all three countries (see Supplementary Table C3). For Extraversion attempts to create a new scale does not represent an improvement over the algorithm-based scale, we refrained from using this new scale for further analyses.

4 Discussion

Previous research has established many meaningful personality-trait-life-outcome associations (Ozer & Benet-Martínez, 2007; Soto, 2019). These results, however, are mainly based on data from Western countries, which represent a small proportion of the world's population. The present study extended research with the Big Six personality traits to three Sub-Saharan African countries. To address the challenges faced when exporting established inventories to different cultural contexts, we created short, measurement-invariant Big Six marker scales suitable for cross-cultural comparisons and tested their suitability and comparability across Kenya, Namibia, and South Africa. We then used these marker scales to explore differences and similarities in life-outcome correlates across the three countries and compared them to what has been established in the West. A major strength of our project was that study materials and methods were developed and contextualized by a collaborative team of scholars from the countries of study (Thalmayer et al., 2024). Our results allow us to draw conclusions on how to best measure the Big Six traits in three African countries, scales which can plausibly be used in other contexts. These findings also provide preliminary insights into which HEXACO/Big Six and related Big Five traits may be more universally applicable and which may be more context specific. Identifying ubiquitous trait aspects is crucial (Leising et al., 2022). If personality researchers worldwide continue to build on identifying these core variables through a joint effort of epistemic iteration (Eronen & Bringmann, 2021), this could help the field move beyond its tendency to uncritically overstate the universality of Western findings (Syed, 2024), leading to a nuanced understanding of the interplay of personality and culture, and a more generalizable science that is truly representative of the world's population.

4.1 Measuring the Big Six Traits with Marker Scales in Sub-Saharan Africa

Although prior studies have highlighted difficulties related to exporting personality models and inventories, especially to majority-world contexts, our results indicate that it is possible and practically useful to identify a set of items representing the best "cross-culturally common denominator" for each Big Six trait. Most of our marker scales fit remarkably well based on standard benchmarks (Hu & Bentler, 1999). This is noteworthy given that personality models tend to have poor CFA fit, even within the populations in which they were developed (Hopwood & Donnellan, 2010), and that model fit is often worse outside Western countries (García et al., 2022; Lajaaj et al., 2019; Thalmayer et al., 2022). Here, measurement invariance tests indicate that the basic organization of the latent variables measured through our marker scales is similar across the three countries (Fischer & Karl, 2019). This does not imply that personality traits naturally organize in a uniform structural model across and within these contexts, as we did not compare models or use an approach that developed local structures from the ground up (e.g., Thalmayer, Rotzinger, et al., 2023). However, the scales we developed form a solid basis for making some types of cross-cultural comparisons with regards to commonly-used personality traits.

The fact that the highest level of invariance reached in our study was metric aligns with prior cross-national studies of personality scales (Dong & Dumas 2020). Our Emotionality and Extraversion marker scales performed slightly worse than results reported for the HEXACO-100 and the HEXACO PI-R across large cross-national samples, where metric invariance was established based on multigroup CFA (Lee et al., 2018; Thielmann et al., 2020). But these studies did not include samples from Africa. Prior studies of HEXACO and Big Six questionnaires such as the HEXACO-60 and the Big Six Questionnaire

including African samples have only achieved configural invariance (García et al., 2022[5]; Thalmayer, Saucier, et al., 2021; Thalmayer & Saucier, 2014). For Emotionality, at least, metric invariance was established in our post-hoc analyses, leading us to conclude with scales for five of the six traits that work as well in and across these African contexts as has been found for the full inventory in contexts less culturally-distant from the Big Six's origin.

Internal consistency did not meet our prespecified cutoff value and was substantially lower than established Big Six scales tested in other contexts (see Ion et al., 2017 and Thielmann et al., 2020 who report moderate to high internal consistencies for the HEXACO-100 and the HEXACO-60 in Asia and Europe). However, in comparison to previous work testing HEXACO/Big Six scales in Sub-Saharan Africa, our marker scales have comparable (García et al., 2022) or better (Thalmayer, Saucier, et al., 2021) internal consistency. Because our study used cross-sectional self-report data for the Big Six, we focused on assessing reliability through measures of internal consistency. Importantly, low internal consistency does not imply that our marker scales are not reliable per see (Ziegler et al. 2014); to obtain a full picture, additional reliability parameters are necessary (e.g., interrater reliability).

Taken together, our findings indicate that our Big Six marker scales may be better suited for assessing common, broad personality traits in Sub-Saharan Africa than other available options, including established Big Five and Big Six/Big Six inventories. Using single term marker items with a standardized stem instead of short phrases avoids some of the challenges that occur when personality inventories are imported (Thalmayer et al., 2022). This approach has promising applicability and potential generalizability to other cultural contexts (Romano et al., 2023).

4.2 More Universal? Honesty, Agreeableness, Conscientiousness, and Openness

Our findings suggest that some Big Six traits may be more cross-culturally transferable than others: for Honesty Agreeableness, Conscientiousness, and Openness we established metric invariance across Kenya, Namibia, and South Africa, indicating that the understanding and expression of these concepts is similar across the three countries as well as to Western contexts. Moreover, associations between Honesty, Agreeableness, Conscientiousness, and Openness with outcome criteria both measured at W1 and W5 were similar to what has been reported in other cultural contexts, indicating that these traits are to some extent functionally universal. Specifically, our findings are in line with Aghababaei and colleagues (2016) who reported positive associations between Honesty, Agreeableness, Conscientiousness, and religiosity (but see also Silvia et al., 2014). And our results confirm earlier work indicating that Conscientiousness is important for physical health (Pletzer et al., 2023; Stephan et al., 2020; Strickhouser et al., 2017), and consistently related to lower levels of overall mental problems (p-factor) and externalizing tendencies across the three countries (Zettler et al., 2020).

Some cultural differences also emerged for these three Big Six traits, putting some limits on their status as functionally universal. First, most of the correlates were only significant in one or two of the three country samples, indicating that the nomological net of each trait may differ across these contexts. Second, incongruent loading patterns of the Honesty item *sly* and the Agreeableness item *irritable* between Namibia and Kenya/South Africa indicate potential differences in the organization of these traits. These items were dropped in the final scales, nevertheless, we now turn to discussing their

---

[5] Note that García et al. (2022) partly applied different criteria to assess measurement invariance. When judged based on the Δ, CFI <.01 criterium which we applied, configural invariance is reached for all scales.

incongruent loadings to gain a more fully understanding of the cross-cultural applicability of the corresponding traits. One explanation could be the high diversity in ethnicity, languages, geographical, and sociopolitical conditions between the contexts, which might be related to personality differences, or differences in the familiarity, valence, or typical usage of these terms. On the other hand, differences could be due to measurement challenges in a context where survey research is even less familiar, especially with regard to less-familiar terms such as *sly* and *irritable.* This is in line with our finding that the differences were mainly for Namibia, which is the most rural and least industrialized among the three countries and had the highest proportion of participants who indicated not knowing some personality words administered in the study. Kenya and South Africa have more similarities in contextual factors such as higher levels of urbanization, industrial development, and possibly closer socio-economic conditions, although Namibia and South Africa share a colonial history and have closely associated economies and national cultures. Given the diversity in the ALLS samples, however, it seems fair to be more surprised by the high congruency of so many terms across the samples and in comparison to the West, than by finding some differences.

To better understand the contradictory and incongruent item loadings, we conducted a small anonymous post hoc survey of ALLS research assistants, knowing that these respondents would be older and significantly more educated than ALLS participants, but come from similar backgrounds. Respondents from Kenya (n = 19), Namibia (n = 10), and South Africa (n = 9) reported whether they know and use these adjectives to describe people and explained the behaviors that would make them describe a person with these adjectives using an open-ended response format. Results indicated that the contradictory loadings of *sly* and *irritable* could have been caused by lack of familiarity and misunderstandings regarding these items. Detailed results are reported in Supplement D.

All in all, our results suggest that Honesty, Openness, and Agreeableness fall into the category of existential universals, with some functionally universal aspects. They encompass aspects of individual variation that are shared across diverse cultural contexts in that most marker items are associated with the same underlying trait in a similar manner, but the role of Agreeableness did not generalize to South Africa. Moreover, the role of Openness has yet to be tested, using a broader set of outcome variables that have been reported to be more closely related to this trait (e.g., creativity, political attitudes, see Zettler et al., 2020).

More Culturally-specific? Emotionality and Extraversion

For Emotionality and Extraversion we found more cross-cultural divergences, suggesting that these traits are understood more differently across the three countries and compared to the West, and may be more sensitive to contextual factors. For example, our marker scales only reached configural invariance (see also García et al., 2022 for similar results indicating that the Extraversion facet of social self-esteem is non-invariant across diverse countries). However, we were able to establish a metric-invariant scale for Emotionality in a post hoc analysis, and this scale demonstrated some of the anticipated associations with mental and physical health, albeit with some differences across countries. These findings suggest that Emotionality may belong with the group above, including both functionally and existentially universal aspects. The expected associations between the initial Emotionality scale and mental health (e.g., Pletzer et al, 2023; Zettler et al., 2020) could not be consistently confirmed in the ALLS samples, with one likely reason being a surprising mixture of item associations. Individuals in the ALLS who scored high on the Emotionality marker scale described themselves as less *brave* and *independent* but, surprisingly, also as less *emotional*, at least in Namibia (the loading direction was the same in Kenya and South Africa, yet the loadings did not reach significance). In our post hoc query to

research assistants, we found that all were familiar with and use the term *emotional*. . A few responses indicated a fully positive understanding of the word, a clue to an interpretation that may have been more common among our younger and less educated participants, and would explain the "backwards" association of this term with its scale.

The Extraversion scale could not be successfully improved, and thus simply appears to be less transferable to these contexts and thus more culture-specific to the West. Previously well-replicated associations between high Extraversion and better mental health (Zettler et al., 2020) and physical health (Stephan et al., 2020, but see also Pletzler et. al, 2023) only held in South Africa in the present study. Incongruent loading patterns of the Extraversion marker scale compared to prior findings align with studies highlighting difficulties in applying this concept outside of Western contexts, as described below. Individuals with high scores on Extraversion in all three contexts described themselves as more *cheerful*, *sociable*, and *talkative*, but, surprisingly, also as more *passive*, which has heretofore been a reverse-scored indicator in the Big Six framework. We also had to exclude *shy* due to incongruent loadings across countries. In our post hoc query to research assistants, most reported familiarity with and use of these words, and open answers indicated the theoretically-expected understanding of low Extraversion. Thus, we did not gain any insights into the unexpected self-descriptions among ALLS participants. Future work with ALLS participants or other younger and less educated participants in these countries will be needed to clarify if these terms (*shy, passive*) are used in meaningfully different ways.

Drawing definitive conclusions about the aspects of Extraversion that are consistently understood across cultural contexts, as opposed to those aspects where manifestations and perceived meanings differ, remains challenging. Our results support the notion that the concept of Extraversion is less applicable in less-individualized, rural contexts, characterized by lower relational mobility (Yuki & Schug, 2012) and fewer interactions with strangers (Thalmayer, Job, et al., 2021). Corroborating this idea, prior research has demonstrated that indigenous languages in Africa, including some spoken in Namibia and Kenya, have few personality-terms related to Extraversion, which indicates that making distinctions in this domain has not been historically relevant in these contexts (Thalmayer et al., 2020; Thalmayer, Job et al., 2021). There is also indicator that the associations are more negative, such as with arrogance, rather than positive (Thalmayer, Saucier, et al., 2020). This leads to doubts about Extraversion being transferable across diverse cultural contexts in the sense of an existential universal (Fontaine, 2011). It is also possible that the terms that describe Extraversion in Western contexts do not correspond to a single coherent trait in other contexts, which would render it non-universal. The trait concept of Extraversion might be specific to Western contexts, likely those where emotional expression, positive emotions, and self-enhancement are highly valued (Kitayama et al., 2023).

4.3 Future Directions, Implications, Limitations

Our study extends the existing literature on personality trait measurement and outcomes by assessing these questions in highly underrepresented contexts that differ in many socio-cultural respects from the Western countries that dominate psychological research (Henrich et al., 2010; Thalmayer, Toscanelli, et al., 2021). Our findings offer both practical and theoretical implications for cross-cultural personality psychology and open several important pathways for future research. In the following sections, we synthesize these contributions and next steps, with acknowledgement of our study's limitations.

The brevity and simplicity of the 27-item inventory that we developed here will be advantageous for repeated ecological assessments and scenarios, involving extensive item batteries, as in large panel and cross-national studies generally (Romano et al., 2023) and in the ALLS specifically. Employing our marker scales in such designs could, for instance, provide valuable insights with respect to cross-cultural differences in day-to-day personality dynamics. One important usage of these marker scales is in further studies in the ALLS samples, for example to explore the long-term stability of traits and to test if established normative personality trajectories, like personality maturation are consistent across contexts, as well as how traits measured at age 18 predict outcomes later in adulthood. While our results demonstrate that some Big Six traits predict meaningful outcomes at a single time point, future studies can consider a wider range of outcome variables to better understand which aspects of the Big Six's nomological net might be more universal and which more context-specific (Zettler et al., 2020). This entails examining how these traits predict impactful life outcomes such as career success, community involvement, and political attitudes over the long term (Ozer & Benet-Martínez, 2006). Ultimately, such research could contribute to a more global understanding of the predictive power of personality traits. It is also important to test if the Big Six-outcome associations found in our study generalize to other majority-world populations, for example, if and where our finding that Emotionality was not related to mental health replicates in other contexts. However, when applying our marker scales, one should bear in mind that measuring personality traits with a consistent short phrase and single terms offers a specific perspective on personality closely related to self-concept, as opposed to items that describe specific behaviors or actions (Back et al., 2009; Romano et al., 2023; Wiedenroth & Leising, 2020).

Uncovering the specific sociocultural factors that explain contextual variations in the expression and consequences of personality traits is another important goal for further research. Our sample was highly diverse, including members of nearly 50 different ethno-linguistic groups. This diversity in terms of ethnicity, languages, geography, and sociodemographic living conditions is extremely valuable, but it also highlights the challenges of using 'country' as a proxy for cultural context. Other cultural groupings, membership in different kinds of churches, and differences in socio-economic status, cut across the three countries. This was not addressed in the current study, and we hope that other teams, especially African scholars, may develop ideas how to define the most meaningful group contrasts in these samples.

In addition to theoretical questions arising from our research, there are important practical implications that need to be considered when exporting personality models and inventories to other places: Ignoring the contextual embeddedness of personality can lead to biased assessments, erroneous conclusions, and ultimately, problematic outcomes, for example, when personality tests are used for employee selection (Anglim et al., 2018). Another practical field is psychological treatment: In Europe and North-America, psychotherapy sometimes aims to change personality aspects that are considered dysfunctional, for example high emotional reactivity (Sauer-Zavala et al., 2017). Since personality traits appear to associate differentially to mental health in these contexts, for example, with Emotionality seen as much as a strength as a vulnerability. Thus, treatments targeting Emotionality can be harmful if exported without careful consideration (Fanon, 1968; Horn, 2020; Lambo, 2010; Tamir et al., 2023). Better comprehending the relationship between personality traits and mental health will be crucial for crafting effective prevention strategies tailored to local contexts.

A major limitation of our marker scales is their limited content validity, that is, our short marker scales might not provide complete construct coverage for the HEXACO, for example, as a whole. This lack of construct coverage is a natural consequence of our focus on cross-culturally generalizable aspects of the HEXACO that resulted in the omission of more culture-specific aspects. In line with our

focus on generalizability, the item pool used to create our marker scales did not provide a fully comprehensive representation of the theoretical HEXACO construct space (Haynes et al., 1995), because it only included cross-culturally ubiquitous Big Six markers. Furthermore, we excluded items with inconsistent loading patterns.

In addition, based on Olaru's (2019) advice for item-selection algorithms, we focused on the optimization criteria most essential for our research questions, that is, we prioritized model fit, measurement invariance and reliability. Schroeders and colleagues (2016) point out that optimizing model fit and internal consistency leads to scales with similar, highly-correlated items, which might come at the cost of content validity. We overcame problems associated with redundant items by ensuring that items were not highly intercorrelated before applying the item-selection algorithm.

To better understand which aspects of the HEXACO/Big Six construct space are covered by our scales and which are not, we assess here how our items correspond to the facets of the HEXACO-100 (Lee & Ashton, 2018). The detailed results are reported in Supplement C, however, we outline one example here, to illustrate how focus on cross-cultural generalizability can come at the cost of lower construct coverage: Our Openness marker scale includes mainly aspects of creativity and intellect, but does not represent unconventionality, aesthetic appreciation, and inquisitiveness. This pattern aligns with Schwaba's and Thalmayer's (2024) argument that being *artistic* or *broadminded* might be particularly relevant in individualistic, culturally lose contexts, which provide more opportunities for niche-picking, for example with respect to selecting one's work, relationships and leisure activities. Relatedly, in more collectivist, culturally tight context, where social norms are stricter and narrower, traits like unconventionality may hold less relevance. This example highlights why certain aspects of the HEXACO/Big Six construct space may be less generalizable, and therefore, focusing on cross-culturally shared aspects might lead to the exclusion of such more culture-specific aspects. We urge other researchers to take this limitation into account when interpreting our findings and using our scales.

In the light of our research goals and the brevity of our marker scales, we consider the lack of construct coverage a reasonable trade-off. In addition, the observed correlations in line with Big Six theory indicate that our marker scales are valid (Cronbach & Mehehl, 1955; Funder & Gardiner, 2024; Revelle & Garner, 2020). The stable Big-Six-outcome associations found in this study are in line with previous findings, indicating that extensively used short scales with limited construct coverage can be valid predictors of outcome variables (see for example Ehrhart al., 2009, and Thalmayer et al., 2011, who report trait-outcome associations supporting the predictive/concurrent validity of the TIPI and BFI-10).

Nevertheless, further research could examine retest reliability along with convergent and discriminant validity. In addition, a fruitful avenue for further research would be to optimize the scales based on other criteria (e.g., construct coverage, balance between positively and negatively keyed items, prediction of outcomes, construct coverage; see Jankowsky et al., 2020; Olaru, 2019; Steger, 2023). In the interest of possible augmentation of our marker scales for improved construct coverage in future usage, we identified further items of potential utility, including the most frequent item combinations that consistently appeared across the 50 item selection runs (see Supplemental Table A8). This provides a curated selection from which researchers may choose if they wish to extend our marker scales.

Other extensions might be bottom-up: The ALLS includes over 100 single-adjective-personality-items administered using the same response format, chosen to assess the Cross-Cultural Big Two

(Thalmayer et al., in press) and the Khoekhoe Personality Inventory (Thalmayer, Saucier, et al., 2021) as well as the Big Six. These items could be pooled to empirically identify locally-relevant scales that best associate with or predict key outcomes over time.

5 Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

6 References

Aghababaei, N., Wasserman, J. A., & Nannini, D. (2012). The religious person revisited: cross-cultural evidence from the Big Six model of personality structure. *Mental Health, Religion & Culture*, *17*(1), 24–29. https://doi.org/10.1080/13674676.2012.737771

Aghababaei, N., Błachnio, A., Arji, A., Chiniforoushan, M., Tekke, M., & Mehrabadi, A. F. (2015). Honesty–Humility and the Big Six structure of Religiosity and Well-Being. *Current Psychology, 35*(3), 421–426. https://doi.org/10.1007/s12144-015-9310-5

Allik, J., & Realo, A. (2017). Universal and specific in the Five Factor Model of personality. In T. A. Widiger (Ed.), *The Oxford handbook of the Five Factor Model* (pp. 173–190). Oxford University Press

Anglim, J., Lievens, F., Everton, L., Grant, S. L., & Marty, A. (2018). Big Six personality predicts counterproductive work behavior and organizational citizenship behavior in low-stakes and job applicant contexts. *Journal of Research in Personality, 77,* 11–20. https://doi.org/10.1016/j.jrp.2018.09.003

Arnett J. J. (2008). The neglected 95%: why American psychology needs to become less American. *The American psychologist*, *63*(7), 602–614. https://doi.org/10.1037/0003-066X.63.7.602

Ashton, M. C., & Lee, K. (2009). The Big Six-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*(4), 340–345. https://doi.org/10.1080/00223890902935878

Ashton, M. C., & Lee, K. (2019). Religiousness and the Big Six personality factors and facets in a large online sample. *Journal of Personality*, *87*, 1103–1118. https://doi.org/10.1111/jopy.12459

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, *86*(2), 356–366. https://doi.org/10.1037/0022-3514.86.2.356

Back, B., Schmuckle, S., & Egloff, E. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology*, *97*(3). https://doi.org/10.1037/a0016229

Behrens, K. Y. (2004). A Multifaceted view of the concept of Amae: reconsidering the indigenous Japanese concept of relatedness. *Human Development*, *47*(1), 1–27. https://doi.org/10.1159/000075366

Benet-Martínez, V. (2007). Cross-cultural personality research: Conceptual and methodological issues. In *Handbook of research methods in personality psychology* (pp. 170–189). The Guilford Press.

Berkessel, J. B., Gebauer, J. E., Joshanloo, M., Bleidorn, W., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2021). National religiosity eases the psychological burden of poverty. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(39), e2103913118. https://doi.org/10.1073/pnas.2103913118

Bleidorn, W., Klimstra, T. A., Denissen, J. J. A., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality maturation around the world: A cross-cultural examination of social-investment theory. *Psychological Science*, *24*(12), 2530–2540. https://doi.org/10.1177/0956797613498396

Boyce, W., Torsheim, T., Currie, C., & Zambon, A. (2006). The Family Affluence Scale as a measure of national wealth: Validation of an adolescent self-report measure. *Social Indicators Research, 78*(3), 473–487. https://doi.org/10.1007/s11205-005-1607-6 Caribé, et al. (2012)

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: stability and change. *Annual Review of Psychology*, *56*, 453–484. https://doi.org/10.1146/annurev.psych.55.090902.141913

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science: A Journal of the Association for Psychological Science*, *2*(2), 119–137. https://doi.org/10.1177/2167702613497473

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Clausen, J. A. (1968). Recent developments in socialization theory and research. *The ANNALS of the American Academy of Political and Social Science*, *377*(1), 139–155. https://doi.org/10.1177/000271626837700115

Olaru, G., Schroeders, U., Hartung, J., & Wilhelm, O. (2019). Ant colony optimization and local weighted structural equation modeling. A tutorial on novel item and person sampling procedures for personality research. *European Journal of Personality*, *33*(3), 400–419. https://doi.org/10.1002/per.2195

Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., Mullins-Sweatt, S. N., Shackman, A. J., Skodol, A. E., South, S. C., Sunderland, M., Waszczuk, M. A., Zald, D. H., Afzali, M. H., Bornovalova, M. A., Carragher, N., Docherty, A. R., Jonas, K. G., Krueger, R. F., … Eaton, N. R. (2019). A hierarchical taxonomy of psychopathology can transform mental health research. *Perspectives on Psychological Science*, *14*(3), 419–436. https://doi.org/10.1177/1745691618810696

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281–302. https://doi.org/10.1037/h0040957

DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general self-rated health question. A meta-analysis. *Journal of general internal medicine*, *21*(3), 267–275. https://doi.org/10.1111/j.1525-1497.2005.00291.x

DeYoung, C. G. (2015). Openness/intellect: A dimension of personality reflecting cognitive exploration. In M. Mikulincer, P. R. Shaver, M. L. Cooper, & R. J. Larsen (Eds.), *APA handbook of personality and social psychology, Vol. 4. Personality processes and individual differences* (pp. 369–399). American Psychological Association. https://doi.org/10.1037/14343-017

Dixon R. M. (1982). *Where have all the adjectives gone?: And other essays in semantics and syntax*. Walter de Gruyter.

Durkee, P. K., Lukaszewski, A. W., von Rueden, C. R., Gurven, M. D., Buss, D. M., & Tucker-Drob, E. M. (2022). Niche Diversity Predicts Personality Structure Across 115 Nations. *Psychological science*, *33*(2), 285–298. https://doi.org/10.1177/09567976211031571

Dong, Y., & Dumas, D. (2020). Are personality measures valid for different populations? A systematic review of measurement invariance across cultures, gender, and age. *Personality and Individual Differences*, *160*, 109956. https://doi.org/10.1016/j.paid.2020.109956

Ehrhart, M. G., Ehrhart, K. H., Roesch, S. C., Chung-Herrera, B. G., Nadler, K., & Bradshaw, K. (2009). Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. *Personality and Individual Differences, 47*(8), 900–905. https://doi.org/10.1016/j.paid.2009.07.012

Eronen, M. I., & Bringmann, L. F. (2021). The Theory Crisis in Psychology: How to Move Forward. Perspectives on Psychological Science, 16(4), 779-788. https://doi.org/10.1177/1745691620970586

Ezirim, G. E., Mbah, P. O., Nwagwu, E. J., Eze, I. C., Nche, G. C., & Chukwuorji, J. C. (2021). Trust and trustworthiness in a sub-saharan african sample: contributions of personality and religiosity. *Social Indicators Research*, *153*(3), 1087–1107. https://doi.org/10.1007/s11205-020-02536-z

Fanon, F. (1968). *The Wretched of the Earth* (First Edition First Printing). Grove Press.

Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, *10*, 1507. https://doi.org/10.3389/fpsyg.2019.01507

Fontaine, J. R. J. (2011). A fourfold conceptual framework for cultural and cross-cultural psychology: Relativism, construct universalism, repertoire universalism and absolutism. In *Fundamental questions in cross-cultural psychology* (pp. 165–189). Cambridge University Press. https://doi.org/10.1017/CBO9780511974090.008

Funder, D. C., & Gardiner, G. (2024). MIsgivings about measurement invariance. European Journal of Personality, 38(6), 889-895. https://doi.org/10.1177/08902070241228338

García, L. F., Aluja, A., Rossier, J., Ostendorf, F., Glicksohn, J., Oumar, B., Bellaj, T., Ruch, W., Wang, W., Kövi, Z., Ścigała, D., Čekrlija, Đ., Stivers, A. W., Di Blas, L., Valdivia, M., Ben Jemaa, S., Atitsogbe, K. A., & Hansenne, M. (2022). Exploring the stability of Big Six-60 structure and the association of gender, age, and social position with personality traits across 18 countries. *Journal of Personality*, *90*(2), 256–276. https://doi.org/10.1111/jopy.12664

Gebauer, J. E., Bleidorn, W., Gosling, S. D., Rentfrow, P. J., Lamb, M. E., & Potter, J. (2014). Cross-cultural variations in Big Five relationships with religiosity: A sociocultural motives perspective. *Journal of Personality and Social Psychology*, *107*, 1064–1091. https://doi.org/10.1037/a0037683

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*(2), 354–370. https://doi.org/10.1037/a0030841

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Heine, S. J. (2010). Cultural psychology. In *Handbook of social psychology, Vol. 2, 5th ed* (pp. 1423–1464). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470561119.socpsy002037

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hofstede, G. (2011). Dimensionalizing cultures: the Hofstede model in context. *Online Readings in Psychology and Culture*, *2*(1). https://doi.org/10.9707/2307-0919.1014

Horn, J. (2020). Decolonising emotional well-being and mental health in development: African feminist innovations. *Gender & Development*, *28*(1), 85–98. https://doi.org/10.1080/13552074.2020.1717177

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

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ion, A., Iliescu, D., Aldhafri, S., Rana, N., Ratanadilok, K., Widyanti, A., & Nedelcea, C. (2017). A cross-cultural analysis of personality structure through the lens of the Big Six model. *Journal of Personality Assessment*, *99*(1), 25–34. https://doi.org/10.1080/00223891.2016.1187155

Jankowsky, K., Olaru, G., & Schroeders, U. (2020). Compiling measurement invariant short scales in cross–cultural personality assessment using ant colony optimization. *European Journal of Personality*, *34*(3), 470–485. https://doi.org/10.1002/per.2260

Jia, R., Bahoo, R., Cai, Z., & Jahan, M. (2022). The Big Six Personality Traits of Higher Achievers at the University Level. *Frontiers in psychology*, *13*, 881491. https://doi.org/10.3389/fpsyg.2022.881491

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., Preacher, K., Coffman, D., Rhemtulla, M., Robitzsch, A., Enders, C., Arslan, R., Clinton, B., Panko, P., Merkle, E., Chesnut, S., … Johnson, A. R. (2022). *semTools: Useful Tools for Structural Equation Modeling* (.00.5-6) [Computer software]. https://cran.r-project.org/web/packages/semTools/index.html

Kasen, S., Wickramaratne, P., Gameroff, M. J., & Weissman, M. M. (2012). Religiosity and resilience in persons at high risk for major depression. *Psychological Medicine*, *42*(3), 509–519. doi:10.1017/S0033291711001516

Kim H., Schimmack U., Oishi S., Tsutsui Y. (2018). Extraversion and life satisfaction: A cross-cultural examination of student and nationally representative samples. *Journal of Personality*, 86(4), 604–618. https://doi.org/10.1111/jopy.12339

Kinyanjui DW, Sum AM (2023) Personality traits and substance use among college students in Eldoret, Kenya. *PLoS ONE 18*(5): e0286160. https://doi.org/10.1371/journal.pone.0286160

Kitayama, S., & Park, J. (2021). Is conscientiousness always associated with better health? A u.s.-japan cross-cultural examination of biological health risk. *Personality & Social Psychology Bulletin*, *47*(3), 486–498. https://doi.org/10.1177/0146167220929824

Kitayama, S., Salvador, C. E., Nanakdewa, K., Rossmaier, A., San Martin, A., & Savani, K. (2022). Varieties of interdependence and the emergence of the Modern West: Toward the globalizing of psychology. *American Psychologist, 77*(9), 991–1006. https://doi.org/10.1037/amp0001073

Koenig, H. G., & Büssing, A. (2010). The Duke University Religion Index (DUREL): a Five-Item measure for use in epidemiological studies. *Religions*, *1*(1), 78–85. https://doi.org/10.3390/rel1010078

Lambo, A. (2010). Psychotherapy in Africa. *Psychotherapy and Psychosomatics*, *24*(4–6), 311–326. https://doi.org/10.1159/000286752

Laajaj, R., Macours, K., Pinzon Hernandez, D. A., Arias, O., Gosling, S. D., Potter, J., Rubio-Codina, M., & Vakis, R. (2019). Challenges to capture the big five personality traits in non-WEIRD populations. *Science Advances*, *5*(7), eaaw5226. https://doi.org/10.1126/sciadv.aaw5226

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the Big Six Personality Inventory. *Multivariate Behavioral Research*, *39*(2), 329–358. https://doi.org/10.1207/s15327906mbr3902_8

Lee, K., & Ashton, M. C. (2006). Further assessment of the Big Six Personality Inventory: two new facet scales and an observer report form. *Psychological assessment*, *18*(2), 182.

https://doi.org/10.1037/1040-3590.18.2.182

Lee, K., & Ashton, M. C. (2008). The Big Six personality factors in the indigenous personality lexicons of English and 11 other languages. *Journal of Personality*, *76*(5), 1001–1054.

https://doi.org/10.1111/j.1467-6494.2008.00512.x

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the Big Six-100. *Assessment*, *25*(5), 543–556. https://doi.org/10.1177/1073191116659134

Lee, K., Ashton, M. C., Griep, Y., & Edmonds, M. (2018). Personality, religion, and politics: An investigation in 33 countries. *European Journal of Personality*, *32*, 100–115. https://doi.org/10.1002/per.2142

Leising, D., Thielmann, I., Glöckner, A., Gärtner, A., & Schönbrodt, F. (2022). Ten Steps Toward a Better Personality Science – How Quality May Be Rewarded More in Research Evaluation. Personality Science, 3(1). https://doi.org/10.5964/ps.6029

McCrae, R. R. (2000). Trait psychology and the revival of personality and culture studies. *American Behavioral Scientist*, *44*(1), 10–31. https://doi.org/10.1177/00027640021956062

McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509–516. https://doi.org/10.1037/0003-066X.52.5.509

Nel, J. A., Adams, B. G., van de Vijver, F. J. R., Laher, S., Louw, J., Makhale, L. M., Naude, L., & Tadi, F. (2017). Personality and well-being in Black and White South African emerging adults. *Current Issues in Personality Psychology, 5*(1), 11–20. https://doi.org/10.5114/cipp.2017.64168

Olaru, G., Schroeders, U., Hartung, J., & Wilhelm, O. (2019). Ant colony optimization and local weighted structural equation modeling. A tutorial on novel item and person sampling procedures for personality research. *European Journal of Personality*, *33*(3), 400–419. https://doi.org/10.1002/per.2195

Oshio, A., Abe, S., Cutrone, P., & Gosling, S. D. (2014). Further validity of the Japanese version of the Ten Item Personality Inventory (TIPI-J): Cross-language evidence for content validity. *Journal of Individual Differences*, *35*(4), 236–244. https://10.1027/1614-0001/a000145

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421. https://doi.org/10.1146/annurev.psych.57.102904.190127

Pew Research Center. (2018, June 13). *The age gap in religion around the world*. https://www.pewforum.org/2018/06/13/the-age-gap-in-religion-around-the-world/

Picard, R. R., & Cook, R. D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association*, *79*(387), 575–583. https://doi.org/10.1080/01621459.1984.10478083

Pletzer, J. L., Thielmann, I., & Zettler, I. (2023). Who is healthier? A meta-analysis of the relations between the Big Six personality domains and health outcomes. *European Journal of Personality*, 089020702311745. https://doi.org/10.1177/08902070231174574

Rammstedt, B., Lechner, C. M., & Danner, D. (2020). Short forms do not fall short. *European Journal of Psychological Assessment*, *37*(1), 23–32. https://doi.org/10.1027/1015-5759/a000574

Revelle, W., & Garner, K. M. (2020). measurement: reliability, construct validation, and scale construction. In *Handbook of Research Methods in Social and Personality Psychology* (2nd ed., pp. 473–503). Cambridge University Press. https://doi.org/10.1017/CBO9780511996481.023

Romano, D., Costantini, G., Richetin, J., & Perugini, M. (2023). The Big Six adjective scales and its psychometric properties. *Assessment*, *30*(8), 2510–2532. https://doi.org/10.1177/10731911231153833

Ronneberg, C. R., Miller, E. A., Dugan, E., & Porell, F. (2016). The protective effects of religiosity on depression: A 2-year prospective study. *The Gerontologist, 56*(3), 421–431. https://doi.org/10.1093/geront/gnu073

Rosseel, Y. (2012). lavaan: An *R* Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Sandy, C. J., Gosling, S. D., & Koelkebeck, T. (2014). Psychometric comparison of automated vs. rational methods of scale abbreviation: An illustration using a brief measure of values. *Journal of Individual Differences, 35*(4), 221–235.doi: 10.1027/1614-0001/a000144

Saucier, G. (2009). Recurrent personality dimensions in inclusive lexical studies: indications for a Big Six structure. *Journal of Personality*, *77*(5), 1577–1614. https://doi.org/10.1111/j.1467-6494.2009.00593.x

Saucier, G., Thalmayer, A. G., Payne, D. L., Carlson, R., Sanogo, L., Ole-Kotikash, L., Church, A. T., Katigbak, M. S., Somer, O., Szarota, P., Szirmák, Z., & Zhou, X. (2014). A basic bivariate structure of personality attributes evident across nine languages. *Journal of Personality*, *82*, 1–14. https://doi.org/10.1111/jopy.12028

Saucier, G., Thalmayer, A. G., & Bel-Bahar, T. S. (2014). Human attribute concepts: Relative ubiquity across twelve mutually isolated languages. *Journal of Personality and Social Psychology, 107*(1), 199–216. https://doi.org/10.1037/a0036492

Sauer-Zavala, S., Wilner, J. G., & Barlow, D. H. (2017). Addressing neuroticism in psychological treatment. *Personality Disorders: Theory, Research, and Treatment*, *8*(3), 191–198. https://doi.org/10.1037/per0000224

Saroglou, V. (2010). Religiousness as a cultural adaptation of basic traits: A five-factor model perspective. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, *14*(1), 108–125. https://doi.org/10.1177/1088868309352322

Schimmack, U., Radhakrishnan, P., Oishi, S., Dzokoto, V., & Ahadi, S. (2002). Culture, personality, and subjective well-being: Integrating process models of life satisfaction. *Journal of Personality and Social Psychology, 82*(4), 582–593. https://doi.org/10.1037/0022-3514.82.4.582

Schmitt, D. P., & Fuller, R. C. (2015). On the varieties of sexual experience: Cross-cultural links between religiosity and human mating strategies. *Psychology of Religion and Spirituality*, *7*, 314–326. https://doi.org/10.1037/rel0000036

Schroeders, U., Wilhelm, O., & Olaru, G. (2016). The influence of item sampling on sex differences in knowledge tests. *Intelligence*, *58*, 22–32. https://doi.org/10.1016/j.intell.2016.06.003

Schwaba, T., & Thalmayer, A. G. (2024). Openness/intellect: The weirdest trait in the Big Five world?. https://doi.org/10.31234/osf.io/wnzy4

Schwartz, S. H. (2007). Value orientations: Measurement, antecedents and consequences across nations. In R. Jowell, C. Roberts, R. Fitzgerald, & G. Eva (Eds.), *Measuring attitudes cross-nationally: Lessons from the European Social Survey* (pp. 169–203). Sage Publications, Inc. https://doi.org/10.4135/9781849209458.n9

Schultze, M., & Schüller, J. (2022). *stuart: Subtests Using Algorithmic Rummaging Techniques* (.00.10.1) [Computer software]. https://cran.r-project.org/web/packages/stuart/index.html

Singh, J. K. (2021). Colonization of personality psychology in India: Historical roots and contemporary status. *Indian Journal of History of Science*, *56*(2), 115–130. https://doi.org/10.1007/s43539-021-00011-3

Silvia, P. J., Nusbaum, E. C., & Beaty, R. E. (2014). Blessed are the meek? Honesty–humility, agreeableness, and the Big Six structure of religious beliefs, motives, and values. *Personality and Individual Differences*, *66*, 19–23. https://doi.org/10.1016/j.paid.2014.02.043

Soto, C. J. (2019). H ow replicable are links between personality traits and consequential life outcomes? the life outcomes of personality replication project. *Psychological Science*, *30*(5), 711–727. https://doi.org/10.1177/0956797619831612

Steger, D., Jankowsky, K., Schroeders, U., & Wilhelm, O. (2023). The road to hell is paved with good intentions: How common practices in scale construction hurt validity. *Assessment, 30*(6), 1811–1824. https://doi.org/10.1177/10731911221124846

Strickhouser, J. E., Zell, E., & Krizan, Z. (2017). Does personality predict health and well-being? A metasynthesis. *Health Psychology*, *36*, 797–810. https://doi.org/10.1037/hea0000475Tamir, M., Ito, A., Miyamoto, Y., Chentsova-Dutton, Y., Choi, J. H., Cieciuch, J., Riediger, M., Rauers, A., Padun, M., Kim, M. Y., Solak, N., Qiu, J., Wang, X., Alvarez-Risco, A., Hanoch, Y., Uchida, Y., Torres, C., Nascimento, T. G., Afshar Jahanshahi, A., . . . García Ibarra, V. J. (2023). Emotion regulation strategies and psychological health across cultures. *American Psychologist.* Advance online publication. https://doi.org/10.1037/amp0001237

Syed, M. (2024). Where are race, ethnicity, and culture in personality research? *Personality Science*, *5*. https://doi.org/10.1177/27000710241257348

Tarakeshwar, N., Stanton, J., & Pargament, K. I. (2003). Religion: An overlooked dimension in cross-cultural psychology. *Journal of Cross-Cultural Psychology*, *34*(4), 377–394. https://doi.org/10.1177/0022022103034004001

Terracciano, A., Löckenhoff, C. E., Crum, R. M., Bienvenu, O. J., & Costa, P. T. (2008). Five-Factor model personality profiles of drug users. *BMC Psychiatry*, *8*(1), 22. https://doi.org/10.1186/1471-244X-8-22

Thalmayer, A.G., Asatsa, S., Shino, E. N., Naudé, L., Laher, S., Hassem, T., Florence, M., Adonis, T.A., Uugwanga, S., Rotzinger, J. S., Hofmann, D., Makunda, J., Botha, C., Murangi, A., and Shirima, C. M. (2024). Lifespan Research in Kenya, Namibia, and South Africa: Cohort Profile of the Africa Long Life Study. *Personality Science*. https://doi.org/10.1177/27000710241264492

Thalmayer, A. G., Job, S., Shino, E. N., Robinson, S. L., & Saucier, G. (2021). ǂŪsigu: A mixed-method lexical study of character description in Khoekhoegowab. *Journal of Personality and Social Psychology*, *121*(6), 1258–1283. https://doi.org/10.1037/pspp0000372

Thalmayer, A. G., Saucier, G., & Eigenhuis, A. (2011). Comparative validity of Brief to Medium-Length Big Five and Big Six Personality Questionnaires. *Psychological Assessment, 23*(4), 995–1009. https://doi.org/10.1037/a0024165

Thalmayer, A. G., Maher, K. A., Saucier, G., Naudé, L., Shino, E. N., Asatsa, S., Florence, M., Witzlack-Makarevich, A., Bächlin, L. M., & Condon, D. (in press) The Cross-Cultural Big Two: A Culturally De-Centered Theoretical and Measurement Model for Personality Traits. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000528 Thalmayer, A. G., Saucier, G., Shino, E. N., & Job, S. (2021). The Khoekhoegowab Personality Inventory: The Comparative Validity of a Locally Derived Measure of Traits. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.694205

Thalmayer, A.G., Marshall, J., Scalise, K. (2023). The International Mental Health Assessment: Validation of an efficient screening inventory. *Collabra: Psychology, 9*(1), 74546.

Thalmayer, A. G., Saucier, G., Ole-Kotikash, L., & Payne, D. (2020). Personality structure in east and west Africa: Lexical studies of personality in Maa and Supyire-Senufo. *Journal of personality and social psychology*, *119*(5), 1132–1152. https://doi.org/10.1037/pspp0000264

Thalmayer, A., Toscanelli, C., & Arnett, J. (2021). Neglected 95% revisited. *American Psychologist*, *76*, 116–129.

Thalmayer, A. G., Saucier, G., & Rotzinger, J. S. (2022). Absolutism, relativism, and universalism in personality traits across cultures: the case of the Big Five. *Journal of Cross-Cultural Psychology*, *53*(7–8), 935–956. https://doi.org/10.1177/00220221221111813

Thielmann, I., Akrami, N., Babarović, T., Belloch, A., Bergh, R., Chirumbolo, A., Čolović, P., de Vries, R. E., Dostál, D., Egorova, M., Gnisci, A., Heydasch, T., Hilbig, B. E., Hsu, K.-Y., Izdebski, P., Leone, L., Marcus, B., Međedović, J., Nagy, J., … Lee, K. (2020). The Big Six–100 across 16 languages: a large-scale test of measurement invariance. *Journal of Personality Assessment*, *102*(5), 714–726. https://doi.org/10.1080/00223891.2019.1614011

Toscanelli, C., Shino, E., Robinson, S. L., & Thalmayer, A. G. (2022). Religiousness worldwide: translation of the Duke University Religion Index into 20 languages and validation across 27 nations. *Measurement Instruments for the Social Sciences*, *4*(1), 13. https://doi.org/10.1186/s42409-022-00041-2

Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & Lucca, N. (1988). Individualism and collectivism: Cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology*, *54*, 323–338. https://doi.org/10.1037/0022-3514.54.2.323

Vygotsky, L. (2020). Educational Psychology. In *CRC Press eBooks*. https://doi.org/10.4324/9780429273070

Wiedenroth, A., & Leising, D. (2020). What's in an adjective?: an exploratory study of meaning systems. *Journal of Individual Differences*, *41*(3), 152–159. https://doi.org/10.1027/1614-0001/a000316

Willroth, E. C., & Atherton, O. E. (2024). Best laid plans: A guide to reporting preregistration deviations. *Advances in Methods and Practices in Psychological Science, 7*(1), Article 25152459231213802. https://doi.org/10.1177/25152459231213802

Yuki, M., & Schug, J. (2012). Relational mobility: A socio-ecological approach to personal relationships. In O. Gillath, G. Adams, & A.D. Kunkel (Eds.), *Relationship Science: Integrating Evolutionary, Neuroscience, and Sociocultural Approaches* (pp. 137–152). https://dx.doi.org/10.1037/13489-007

Zettler, I., Thielmann, I., Hilbig, B., & Moshagen, M. (2020). The nomological net of the Big Six model of personality: a large-scale meta-analytic investigation. *Perspectives on Psychological Science*, *15*, 723–760. https://doi.org/10.1177/1745691619895036

Ziegler, M., Kemper, C. J., & Kruyen, P. (2014). Short scales – Five misunderstandings and ways to overcome them [Editorial]. *Journal of Individual Differences, 35*(4), 185–189. https://doi.org/10.1027/1614-0001/a000148

- We created short, single-term Big 6 scales using an item selection algorithm
- The scales were measurement invariant across Kenya, Namibia, and South Africa
- The Big Six were robustly associated with mental/physical health and religiosity
- Honesty, Agreeableness, Conscientiousness and Openness matched US/European findings

- Extraversion and Emotionality showed cross-cultural differences

Table 1

Hypotheses: Expected and Observed Associations between Big Six Traits and Mental Health, Physical Health, and Religiosity

Variable	Hypotheses	Observed Associations		
		Kenya	Namibia	South Africa
Mental Health p-factor	-		a	a
Life Stress	* *		a	a
Interpersonal Conflict	* *		a	b
Partner Conflict				

nternalizing Spectrum						a	a
nxiety	*	*			a	a	b
epression	*	*			a	a	
xternalizing Spectrum		*	*	a	a	a	a
ubstance (ab)use	*	*	*			a	a
nger		*			a	a	

hysical Health (GSRH)	*	*	*		a
eligious Activity (DUREL)	*		*	*	a
ntrinsic Religiosity (DUREL)	*		*	*	a

Note. Hypotheses are based on theory and previous evidence from mainly Western samples. + in the Hypotheses column indicates that we expected a positive correlation, - indicates that we expected a negative association. * indicates that a hypothesis was confirmed in at least one country, that is, the correlation was in the expected direction and significant at $p < .01$. + in the Observed Associations column represent a significant positive bivariate correlation ($p < .01$) in our data, - in the Observed Associations column represent a significant negative bivariate correlation ($p < .01$). [a] indicates that the association was also significant at $p < .01$ in a regression analysis with all Big Six traits as predictors. [b] indicates that the association was only significant in a regression analyses with all Big Six traits as predictors.

Note that the direction of the hypotheses for the Mental Health p-factor were mistakenly reversed in our preregistration. This correction aligns with the correctly preregistered hypotheses for the mental health subscales (e.g., internalizing and externalizing spectra, and life stress) from which the p-factor is derived. In addition, we did not preregister separate hypotheses for religious activity and

intrinsic religiosity because we initially planned to use the DUREL Religiosity index as an overall indicator of religiosity.

Table 2

Sociodemographic Characteristics of the Analytic Sample

Characteristics	Kenya	Namibia	South Africa
	%	%	%
Gender			
Female	46	55	74
Male	54	45	25
Other	0	0	0
NA	0	0	0
Language Family			
Bantu	66	65	72
Khoisan	0	30	0
Nilotic	28	0	0
Afro-Asiatic	4	0	0
Indo-European	0	5	28
NA	1	0	0

Note: Due to rounding, the sum of percentages does not equal 100% in some cases.

Table 3

Big Six Marker Items included in the Africa Long Life Study

Honesty/Humility (7)	Emotionality (8)	Extraversion (11)
boast	anxious	**cheerful***
greedy	**brave***	introverted
honest*	courageous*	lively*
just	**emotional**	**passive**
loyal	**independent**	quiet*
sincere	**self-assured**	reserved
sly	**strong**	shy*
	vulnerable	silent*
		sociable
		solitary
		talkative

Agreeableness (10)	Conscientiousness (11)	Openness (8)
stubborn	**careless**	**clever***
aggressive*	**diligent***	**creative**
agreeable	**disciplined**	**gifted**
good-natured*	industrious*	intellectual
hot-headed*	**irresponsible**	**intelligent**
irritable	lazy	ironic
mild	orderly	original

patient	organized	**sharp**
peaceful*	precise	
tolerant	**reckless**	
	thorough	

Note. Items with * were administered at W1, all others were administered at W2. Items in bold were included in the final Big Six marker scales chosen through the present analyses.

Table 4

Internal Consistencies, Means, and Standard Deviations

Scale	Combined Sample				Kenya				Namibia				South Africa			
		D				*D*				*D*				*D*		
Big Six Marker Scales																
Honesty/Humility	.97	.90	55	62	.25	.73	59	68	.65	.06	50	53	.06	.68	39	59
Emotionality	.36	.83	45	60	.27	.73	33	52	.48	.94	53	65	.30	.76	44	66
Extraversion	.30	.97	40	50	.50	.90	39	51	.04	.01	27	31	.42	.92	52	65
Agreeableness	.70	.90	48	53	.98	.80	43	50	.43	.00	46	54	.72	.77	44	49
Conscientiousness	.85	.96	61	67	.24	.77	59	67	.65	.98	47	52	.64	.97	65	73
Openness	.80	.87	66	70	.12	.72	63	68	.57	.98	65	69	.72	.77	63	67

Mental
ental

Health
(IMHA)

p-factor	.11	.84	92	93	.95	.78	92	93	.22	.90	92	93	.14	.79	92	93
Life Stress	.73	.28	65	70	.56	.21	65	67	.66	.30	70	71	.02	.29	65	74
Interpersonal Conflict	.44	.38	62	65	.27	.28	62	63	.63	.41	58	61	.40	.41	65	72
Partner Conflict	.54	.02	80	83	.53	.99	80	89	.69	.09	76	78	.37	.95	86	89
Internalizing Spectrum	.51	.14	92	92	.3	.07	92	93	.47	.05	90	91	.83	.25	92	93
Anxiety	.58	.28	79	82	.35	.2	79	82	.52	.19	76	80	.93	.41	82	84
Depression	.63	.31	84	85	.32	.22	75	87	.57	.18	72	79	.06	.45	74	87
Externalizing Spectrum	.70	.82	82	88	.6	.69	85	87	.97	.97	79	89	.46	.64	86	88
Substance (ab)use	.42	.93	86	88	.25	.73	79	87	.62	.12	83	88	.35	.82	78	88
Anger	.99	.08	76	79	.94	.98	84	77	.33	.23	87	79	.58	.76	83	74

Physical Health	.75	.04			.65	.03			.81	.06			.79	.02		
DUREL																
Religious Activity (DUREL)	.93	.77	66	67	.22	.35	68	70	.21	.79	58	60	.39	.68	63	65
Intrinsic Religiosity (DUREL)	0.97	.27	67	72	1.49	.16	65	74	0.05	.24	60	62	1.61	.17	76	81
Socioeconomic Status (FAS)	.80	.22	53	86	.31	.16	60	74	.58	.96	43	54	.65	.38	58	69

Note. α and ω represent Cronbach's Alpha and McDonald's Omega. Big Six scales refined in this study (excluding items in italics Table 5) are used. For Big Six and mental health scales, mean values of items are used. For Religious Activity, Intrinsic Religiosity, and socio-economic status, score was derived by summing items, leading to possible ranges of 2 – 12, 3 – 15, and 0-9, respectively.

Table 5

Initial Big Six Marker Scales Resulting From the Item Selection Algorithm

Scale	Selected Marker Items	Number of runs in which the item set was selected
H	honest, just, loyal, sincere, *sly*	28

| E | brave, emotional, independent, self-assured, strong | | | | | 16 |

| X | cheerful, passive, *shy*, sociable, talkative | | | | | 7 |

| A | agreeable, good-natured, *irritable*, patient, tolerant | | | | | 13 |

| C | careless, diligent, disciplined, irresponsible, reckless | | | | | 6 |

| O | clever, creative, gifted, intelligent, sharp | | | | | 6 |

Note. We applied the item selection algorithm 50 times in randomly selected halves of the sample and chose the item combination which emerged most frequently across the 50 runs. Items in italics were removed due to inconsistent loading patterns across the three samples.

Table 6

Measurement Invariance of the Final Big Six Marker Scales

Trait		df	χ^2	$p(\chi^2)$	RMSEA	CFI	CFI Δ
Honesty/Humility							
	Configural	6	2.800	.834	.000	1.00	
	Metric	12	11.713	.469	.000	1.00	.000
	Scalar	18	82.657	<.001	.088	.837	.163

Emotionality

Configural	15	21.931	.110	.030	.985	
Metric	23	39.732	.016	.038	.963	.022
Scalar	31	144.12	<.001	.084	.750	.213

Extraversion

Configural	6	12.351	.055	.048	.974	
Metric	12	24.227	.019	.047	.951	.024
Scalar	18	123.223	<.001	.113	.576	.375

Agreeableness

Configural	6	9.814	.133	.036	.984	
Metric	12	13.198	.355	.014	.995	+.011
Scalar	18	64.484	<.001	.072	.810	.185

Conscientiousness

Configural	15	39.253	.001	.061	.960	

Metric	23	46.406	.003	.049	.961	+.001
Scalar	31	121.848	<.001	.083	.850	.111

Openness

al	Configur	15	48.251	<.001	.064	.965	
	Metric	23	54.241	<.001	.050	.967	+.002
	Scalar	31	91.766	<.001	.061	.936	.031

Table 7

Factor Loadings (λ) for the Final Big Six Marker Scales, 95% Confidence Intervals, Standard Errors (SE), and Probability Values per Country

Big Six items	Kenya				Namibia				South Africa			
		5 % CI	E			5 % CI	E			5% CI	E	
Honesty/Humility												
honest	.41	0.27, 0.55]	.07	.001	.60	0.34, 0.85]	.13	.001	.73	0.44, 1.03]	.15	.001
just	.13	0.96, 1.29]	.09	.001	.94	0.67, 1.21]	.14	.001	.77	0.43, 1.11]	.17	.001
loyal	.05	0.89, 1.2]	.08	.001	.13	0.83, 1.43]	.15	.001	.81	0.52, 1.1]	.15	.001
sincere	.41	1.23, 1.6]	.09	.001	.33	0.99, 1.67]	.17	.001	.69	1.15, 2.23]	.28	.001
Emotionality (R)												
brave	.72	0.43, 10]	.14	001	.34	0.11, 0.57]	12	004	.70	0.45, 0.95]	.13	001
emotional	.14	-0.41, 0.70]	.28	62	.62	0.36, 0.87]	13	001	.35	0.01, 0.68]	.17	042
independent	.21	0.83, 1.59]	.19	.001	.39	1.15, 1.63]	.12	.001	.06	0.8, 1.31]	.13	.001

self-assured	.56	1.16, 1.95]	.20	.001	.21	1, 1.43]	.11	.001	.21	0.95, 1.48]	.14	.001
strong	.38	0.98, 1.77]	.20	.001	.44	1.18, 1.7]	.13	.001	.68	1.31, 2.05]	.19	.001

Extraversion

cheerful	.46	0.26, 0.67]	.10	.001	.25	-0.22, 0.72]	.24	303	.78	0.6, 0.96]	.09	.001
passive	.65	0.33, 0.97]	.16	.001	.14	0.44, 1.83]	.36	001	.35	0.13, 0.56]	.11	001
sociable	.07	0.71, 1.44]	.19	.001	.13	0.42, 1.83]	.36	002	.46	1.19, 1.74]	.14	.001
talkative	.81	1.3, 2.32]	.26	.001	.49	0.61, 2.37]	.45	.001	.41	1.15, 1.67]	.13	.001

Agreeableness

agreeable	.29	0.95, 1.62]	.17	001	.47	1.09, 1.85]	.19	001	.22	0.87, 1.58]	18	001
good-natured	.26	0.08, 0.44]	.09	004	.49	0.25, 0.73]	.12	.001	.37	0.14, 0.61]	.12	002
patient	.35	0.99, 1.7]	.18	.001	.11	0.81, 1.41]	.15	.001	.43	1.04, 1.81]	.2	.001
tolerant	.10	0.8, 1.4]	.15	.001	.93	0.65, 1.21]	.14	.001	.98	0.68, 1.28]	.15	.001

Conscientiousness(R)												
careless	.19	1.83, 2.56]	.19	001	.28	1.53, 3.03]	.38	001	.30	1.98, 2.62]	16	001
diligent	0.61	-0.92, -0.29]	.16	.001	0.29	-0.88, 0.3]	.30	335	0.36	-0.62, -0.1]	.13	007
disciplined	1.08	-1.48, -0.68]	.2	.001	1.39	-2.3, -0.48]	.46	003	0.84	-1.16, -0.51]	.17	.001
irresponsible	.27	1.89, 2.64]	.19	.001	.58	1.77, 3.39]	.41	.001	.86	1.58, 2.14]	.14	.001
reckless	.23	1.86, 2.6]	.19	.001	.82	1.17, 2.47]	.33	.001	.04	1.74, 2.33]	.15	.001
Openness												
clever	.44	0.29, 0.6]	.08	001	.41	0.25, 0.57]	.08	001	.56	0.39, 0.74]	09	001
creative	.19	1.02, 1.36]	.09	.001	.07	0.92, 1.22]	.08	.001	.24	1.04, 1.44]	.1	.001
gifted	.13	0.96, 1.31]	.09	.001	.14	0.98, 1.29]	.08	.001	.09	0.89, 1.28]	.1	.001
intelligent	.10	0.93, 1.27]	.09	.001	.12	0.97, 1.27]	.08	.001	.07	0.91, 1.24]	.09	.001

sharp												
	.13	0.97, 1.29]	.08	.001	.26	1.1, 1.42]	.08	.001	.04	0.85, 1.22]	.1	.001

Note. Loadings and intercepts were estimated freely across countries. To identify our models, we applied effects coding to the item loadings, i.e., we constrained the factor loadings of the items for each Big Six trait to average one (Card & Little, 2007)**.** This led to reversed item loadings for Emotionality and Conscientiousness because these marker scales consist of mainly reverse-keyed items. As indicated by the (R) next to Emotionality and Conscientiousness, high item loadings represent low trait levels

Table 8

Final Big Six Measurement Models: Fit per Country

Trait	Kenya		Namibia		South Africa	
	CFI	RMSEA	CFI	RMSEA	CFI	RMSEA
Honesty/Humility	1.000	.000	1.000	.000	1.000	.000
Emotionality	.954	.048	.997	.016	1.000	.000
Extraversion	.982	.036	1.000	.000	.966	.077
Agreeableness	.988	.032	.974	.049	.994	.022
Conscientiousness	.976	.049	.861	.079	.975	.057
Openness	.978	.050	.960	.072	.956	.069

Note. Model fit and reliability indices were computed based on the configural model, where loadings and intercepts were estimated freely across the three countries.

Table 9

Bivariate Correlations between the Final Big Six Marker Scales with W1 Mental Health, Physical Health, and Religiosity

	Honesty			Emotionality			Extraversion		
	Kenya	Namibia	South Africa	Kenya	Namibia	South Africa	Kenya	Namibia	South Africa
IMHA p-factor	-.18 *	-.11 *	-.05	.02	.06	-.02	-.02	-.01	-.06
Life Stress	-.04	.00	.03	-.02	-.03	-.04	-.02	.03	-.12 *
Interpersonal Conflict	.16 *	-.07	-.02	.02	.00	-.04	-.00	.00	-.01
Partner Conflict	-.18	-.16	-.04	-.10	.16	-.05	-.00	-.05	.05
Internalizing Spectrum	.12*	-.05	-.07	.00	.01	.02	.02	-.01	-.17 *
Anxiety	-.05	-.05	.09	-.02	.00	-.02	-.01	.01	-.14 *
Depression	.21*	-.05	-.04	.06	.01	.10	.06	-.01	-.21 *
Externalizing Spectrum	.21 *	-.16 *	-.00	.04	.10	-.10	-.02	-.03	-18 *
Substance (ab)use	.14 *	-.12 *	-.03	.01	.07	-.10	-.00	-.02	-16 *

Anger	-.19 *	-.14 *	-.03	.05	.10	.07	-.03	-.03	.13 *
Physical Health	.09	.06	.08	-.07	-.04	-.12 *	.08	.01	.18 *
Religious Activity	.23 *	.04	.07	.07	-.02	-.04	-.04	.01	.13 *
Intrinsic Religiosity	.09	.11 *	.08	-.08	-.13 *	-.08	.01	.12 *	.17 *

	Agreeableness			Conscientiousness			Openness		
	Kenya	Namibia	South Africa	Kenya	Namibia	South Africa	Kenya	Namibia	South Africa
IMHA p-factor	-.11 *	-.09	.04	-.15 *	-.16 *	-.18 *	-.15 *	-.09	-.06
Life Stress	.01	.02	.02	.01	-.12 *	-.07	-.02	-.00	-.10
Interpersonal Conflict	-.11 *	-.01	.01	-.12 *	-.15 *	-.09	-.11 *	-.07	-.10
Partner Conflict	-.10	-.06	.04	-.19	-.09	-.08	.13	-.15	.18
Internalizing Spectrum	-.06	-.02	.04	-.10	-.11 *	-.15 *	-.14 *	-.04	-.13*
Anxiety	-.01	-.04	.06	-.06	-.13 *	-.11	-.10	-.02	-.11
Depression	-.11 *	-.03	-.01	-.16 *	-.08	-.19 *	-.18 *	-.06	-.18 *

Externalizing Spectrum	-.17 *	-.14 *	.01	-.17 *	-.18 *	-.14 *	-.14 *	-.12 *	.10
Substance (ab)use	-.13 *	-.10 *	.05	-.10	-.12 *	-.13 *	-.12 *	-.09	.11
Anger	-.14 *	-.13 *	.04	-.16 *	-.17 *	-.10*	-.11 *	-.11 *	.06
Physical Health	.09	.08	.11	.02	.12 *	.11	.15*	.08	.16 *
Religious Activity	.10	.00	.12 *	.13 *	.03	.13 *	.19 *	.05	.05
Intrinsic Religiosity	.08	.13 *	.08	.12 *	.09	.15 *	.07	.10	.12 *

*Note. r*s represent bivariate correlations between the mean scores of the Big Six marker scales with the mean scores of the outcome variables. Correlates of Emotionality and Extraversion should not be compared between countries because these scales are not metrically invariant. *p < .01. Exact *p*-values and 95% Confidence Intervals are displayed in the Supplemental Table A5).