Scholar Commons
USF University of South Florida

University of South Florida
Digital Commons @ University of South Florida

USF Tampa Graduate Theses and Dissertations
USF Graduate Theses and Dissertations

March 2022

Creating a Short, Public-Domain Version of the CPAI-2: Using an Algorithmic Approach to Develop Public-Domain Measures of Indigenous Personality Traits

Mukhunth Raghavan
University of South Florida

Follow this and additional works at: https://digitalcommons.usf.edu/etd

 Part of the Business Administration, Management, and Operations Commons, Personality and Social Contexts Commons, and the Quantitative Psychology Commons

Scholar Commons Citation

Raghavan, Mukhunth, "Creating a Short, Public-Domain Version of the CPAI-2: Using an Algorithmic Approach to Develop Public-Domain Measures of Indigenous Personality Traits" (2022). *USF Tampa Graduate Theses and Dissertations.*
https://digitalcommons.usf.edu/etd/9441

This Thesis is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact scholarcommons@usf.edu.

Creating a Short, Public-Domain Version of the CPAI-2: Using an Algorithmic Approach to Develop Public-Domain Measures of Indigenous Personality Traits

by

Mukhunth Raghavan

A thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Arts in Industrial-Organisational Psychology
Department of Psychology
College of Arts and Sciences
University of South Florida

Major Professor: Brenton M. Wiernik, Ph.D.
Stephen Stark, Ph.D.
Marina A. Bornovalova, Ph.D.

Date of Approval:
December 14, 2021

Keywords: Big Five, China, Cross-cultural, psychometrics, item response theory, genetic algorithm.

Copyright © 2022, Mukhunth Raghavan

Table of Contents

List of Tables iii

List of Figures iv

Abstract 1

Chapter 1: Introduction 2
- The Hierarchical Big Five Model 3
- Cross-Cultural Personality Research 4
- The Cross-Cultural (Chinese) Personality Assessment Inventory (CPAI) 6
  - Development and structure. 6
  - Convergence with the Big Five. 8
  - Validity evidence 9
  - Methodological limitations. 10
  - Limitations of proprietary scales. 11
- The Current Study 12
  - Challenges in developing short-form measures. 13

Chapter 2: Method 16
- Phase 1: Analogue Scales Construction Process 16
  - Etic scales. 16
  - Emic scales. 18
- Phase 2: New Emic Scales Validation Process 19
  - Participants. 19
    - Sample Size Planning. 20
  - Measures. 22
    - CPAI-2. 22
    - BFI-2. 23
    - Public domain CPAI-2 analogue scales 23
    - Criterion-related validity. 23
    - Organizational citizenship behavior. 24
    - Counterproductive work behavior. 24
    - Psychological distress 24
  - Data Collection. 25
- Data Analysis 26
- Evaluating Items for the Public-Domain Scales 26
  - Item factor analysis. 27
  - Genetic algorithm. 30
- Evaluating Short-Form Inventories 33

Convergence with full-length scales and structural validity. 33
Correlations with the Big Five. 34
Criterion-related validity. 34
Reliability. 34

Chapter 3: Results 36
Scale Characteristics of all new CPAI items 36
Convergence with full-length scales. 36
Correlations with the Big Five. 37
Criterion-related validity. 37
Shortening the Scales 38
Item Factor Analysis. 38
Convergence with full-length scales. 38
Correlations with the Big Five. 39
Criterion-related validity. 39
Genetic Algorithm. 39
Convergence with full-length scales. 40
Correlations with the Big Five. 41
Criterion-related validity. 41

Chapter 4: Discussion 42
Item Screening Processes 42
Practical Implications of the New CPAI Scales 45
Limitations 49

Chapter 5: Conclusion 52

References 53

Appendix A: Tables 71

Appendix B: Figures 110

Appendix C: Item writing worksheet 136

List of Tables

Table 1. Four factors and 28 scales from the CPAI-2. 71

Table 2. Patterns of correlations of CPAI-2 etic scales with NEO PI-R scales. 78

Table 3. Public domain analogues identified for the CPAI-2 etic scales. 79

Table 4. Mean convergent correlations from the Genetic Algorithm simulation at different sample sizes. 83

Table 5. Sample characteristics. 84

Table 6. Scale means, standard deviations, and Cronbach's alpha for all measures in the overall sample, as well as the Chinese and US samples. 85

Table 7. New CPAI item characteristics for all items in the full sample. 88

Table 8. Correlations among the new CPAI scales and the equivalent CPAI-2 emic scales. 95

Table 9. Correlations between BFI factors and the new CPAI and CPAI-2 scales. 96

Table 10. Correlations between criterion scales and the new CPAI and CPAI-2 scales. 97

Table 11. Item characteristics for items selected using the IFA approach. 98

Table 12. Correlations among the IFA-shortened new CPAI scales and the equivalent CPAI-2 emic scales. 103

Table 13. Correlations between BFI factors and the IFA-shortened new CPAI scales. 103

Table 14. Correlations between criterion scales and the IFA-shortened new CPAI scales. 104

Table 15. Cross-validation correlations between GA-shortened scales and CPAI-2 scale scores at a series of item costs. 104

Table 16. Number of items in retained from each of the GA-shortened scales scales at resulting from a series of item costs. 105

Table 17. Item characteristics for items selected using the GA approach. 106

Table 18. Correlations among the GA-shortened new CPAI scales and the equivalent CPAI-2 emic scales. 108

Table 19. Correlations between BFI factors and the GA-shortened new CPAI scales. 108

Table 20. Correlations between criterion scales and the GA-shortened new CPAI scales. 109

Table 21. Items retained by the IFA and GA approaches. 109

List of Figures

Figure 1. Distribution of New CPAI scale means among Chinese and US samples. 110

Figure 2. Correlations and intercorrelations among New CPAI and CPAI-2 scales.....111

Figure 3. Correlations between New CPAI scales and equivalent CPAI-2 factors.....112

Figure 4. Distributions of correlations between BFI factors and New CPAI and CPAI-2 scales. 113

Figure 5. Correlations between BFI factors and all New CPAI and CPAI-2 scales. 114

Figure 6. Distributions of correlations between criterion scales and New CPAI and CPAI-2 scales. 115

Figure 7. Correlations between criterion scales and all New CPAI and CPAI-2 scales. 116

Figure 8. Scale information and item trace plots for the FAC scale. 117

Figure 9. Scale information and item trace plots for the FAM scale. 117

Figure 10. Scale information and item trace plots for the DEF scale. 117

Figure 11. Scale information and item trace plots for the GM scale. 118

Figure 12. Scale information and item trace plots for the SS scale.....118

Figure 13. Scale information and item trace plots for the TM scale. 119

Figure 14. Scale information and item trace plots for the REN scale. 119

Figure 15. Scale information and item trace plots for the SOC scale. 120

Figure 16. Scale information and item trace plots for the DIS scale.....120

Figure 17. Scale information and item trace plots for the HAR scale. 121

Figure 18. Scale information and item trace plots for the TE scale. 121

Figure 19. Comparing correlations between New CPAI scales and equivalent CPAI-2 factors using the full scale and the two short scales.....122

Figure 20. Comparing distributions of correlations between BFI factors and New CPAI and CPAI-2 scales using the full scale and the two short scales.....123

Figure 21. Comparing correlations between BFI factors and New CPAI and CPAI-2 scales using the full scale and the two short scales.....124

Figure 22. Comparing distributions of correlations between criterion scales and New CPAI and CPAI-2 scales using the full scale and the two short scales.....125

Figure 23. Comparing correlations between criterion scales and all New CPAI and CPAI-2 scales using the full scale and the two short scales.....126

Figure 24. Scale information and item trace plots for FAC items selected using the IFA approach.....127

Figure 25. Scale information and item trace plots for FAM items selected using the IFA approach.....127

Figure 26. Scale information and item trace plots for DEF items selected using the IFA approach.....128

Figure 27. Scale information and item trace plots for GM items selected using the IFA approach.....128

Figure 28. Scale information and item trace plots for SS items selected using the IFA approach.....129

Figure 29. Scale information and item trace plots for TM items selected using the IFA approach.....129

Figure 30. Scale information and item trace plots for REN items selected using the IFA approach.....130

Figure 31. Scale information and item trace plots for SOC items selected using the IFA approach.....130

Figure 32. Scale information and item trace plots for DIS items selected using the IFA approach.....131

Figure 33. Scale information and item trace plots for HAR items selected using the IFA approach.....131

Figure 34. Scale information and item trace plots for TE items selected using the IFA approach. 132

Figure 35. Scale information and item trace plots for FAC items selected using the GA approach. 132

Figure 36. Scale information and item trace plots for FAM items selected using the GA approach. 133

Figure 37. Scale information and item trace plots for DEF items selected using the GA approach. 133

Figure 38. Scale information and item trace plots for TM items selected using the GA approach. 134

Figure 39. Scale information and item trace plots for HAR items selected using the GA approach. 134

Figure 40. Scale information and item trace plots for TE items selected using the GA approach. 135

Abstract

In this study we aimed to create a short, public-domain analogue of the Cross-Cultural (Chinese) Personality Assessment Inventory (CPAI-2; F. M. Cheung et al., 1996). Emic (culture-specific) traits measured by the CPAI-2 are purportedly specific to the Chinese culture and argued to not be fully captured by the consensus Big Five personality trait taxonomy. Research suggests that CPAI-2 traits may have unique predictive power, especially in non-Western contexts. However, research has been hampered by several limitations of the measure. The inventory is proprietary and long, with 341 items forming 28 scales and four factors. Cross-cultural personality research would benefit from a short, public-domain analogue to the CPAI-2 to permit assessment in a wider range of contexts. Using two analytic approaches—item factor analysis and a genetic algorithm (Yarkoni, 2018)—we developed two short, public-domain measures to assess the 11 emic CPAI-2 scales that have no clear analogues in the current public-domain personality measure library. When examining the resulting measures' factor structure, reliability, and criterion-related validity, we see that both short-form measures adequately replicate the pattern of correlations exhibited by the full-form measure as well as the original CPAI-2. Implications for research using automated scale abbreviation and the cultural specificity hypothesis of personality are discussed.

Chapter 1: Introduction

Personality traits are relatively stable characteristics that describe individual differences in thoughts, feelings, and behaviors (John et al., 2008). Measures of personality traits have been found to be strong predictors of important work and life outcomes, including job performance behaviors, work and life satisfaction, and well-being (Connelly et al., 2018; Ozer & Benet-Martínez, 2006; Roberts et al., 2007; Soto, 2019), leading personality assessment to become a core aspect of many areas of psychological research.

Much recent research in personality has been organized around the consensus Hierarchical Big Five (Five-Factor) Model (John et al., 2008; Ones et al., 2016; Stanek & Ones, 2018). This model posits that personality traits can be well-organized around five principle dimensions (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness), along with a variety of higher-order, lower-order, and compound traits. However, although the Hierarchical Big Five Model is widely accepted, several criticisms have been raised regarding its generalizability. Chief among these is the question of whether this personality structure is applicable across cultures (F. M. Cheung et al., 2011; De Fruyt & Wille, 2013; Stankov, 2011). For example, several researchers have suggested that the specific content of the Big Five dimensions varies across cultures (De Raad et al., 2010), or that models with six principle dimensions may replicate more consistently across cultures (Saucier, 2009). Most seriously, several researchers have suggested that the Big Five and other personality frameworks developed in Western contexts are missing important culture-specific personality traits from non-Western cultures (F. M. Cheung et al., 2011).

As an example of the last point, Cheung and colleagues developed the Cross-Cultural (Chinese) Personality Assessment Inventory (CPAI; F. M. Cheung et al., 1996). The CPAI[1] was developed using an indigenous (emic) approach to capture the personality dimensions most relevant in Chinese culture. The CPAI demonstrates some favorable psychometric properties, and it has seen adoption as a measure of indigenous Chinese personality measure. However, the CPAI is extremely long (341 items), limiting its utility for personality research and application when other constructs must also be administered. In addition, the CPAI is a proprietary scale, which hampers the ability of researchers to deeply probe and investigate the captured constructs (Goldberg et al., 2006). In this study, we address these limitations by constructing public-domain short-form measures of the CPAI indigenous trait constructs. We develop new items to assess the indigenous scales, constructing scales using two distinct analytic approaches (item factor analysis and genetic selection algorithms), and compare the reliability and validity of the measures developed using each approach.

The Hierarchical Big Five Model

The Hierarchical Big Five Model has emerged as a consensus model for personality traits, providing an organizational framework for classifying and interpreting measures (for a review of the history of the Big Five, see John et al., 2008; John & Srivastava, 1999). The Big Five dimensions were initially discovered in studies following the "lexical hypothesis," which posits that most of the socially relevant and salient personality characteristics have become encoded into natural language, such that our vocabulary contains key descriptors of the major personality traits. Factor analyses of trait descriptive adjective ratings, primarily in English and German, found consistent support for five broad factors, termed the "Big Five" by Goldberg (1981) to emphasize their breadth and that they

[1] A revised measure was later developed and named "CPAI-2" (F. M. Cheung, Shu Fai Cheung, et al., 2008). When referencing the CPAI, our discussions and critiques in this study apply generally to the CPAI and the CPAI-2 scales unless we explicitly refer to the original CPAI or to the CPAI-2.

each encompass many more specific traits. Concurrently, researchers conducting factor analyses of existing questionnaires developed from a variety of earlier personality theories also found support for five similar principle dimensions (McCrae & Costa, 1997). Subsequently, the hierarchical Big Five structure has received support across a variety of measures, rating sources, and contexts (McCrae & Costa, 1997; Ones et al., 1996), as well as from computational, neuroscientific, and behavioral genetic personality research (DeYoung, 2010, 2015; Yamagata et al., 2006).

Two structural features are key to the Hierarchical Big Five Model (Markon et al., 2005). First, as the name implies, the model specifies that personality traits are hierarchically organized. Each of the Big Five encompasses a number of narrow aspect and facet traits. Further, above the Big Five, very broad "metatraits" capture broad tendencies in behavioral exploration and inhibition (DeYoung, 2015; DeYoung & Krueger, 2018). Traits at each of these hierarchical levels may have different developmental trajectories and patterns of relationships with behavioral and outcome variables. Second, personality traits lack simple structure. Personality items and lower-order traits tend not to load cleanly onto only one of the Big Five factors. Rather, most items and narrow traits show substantial cross-loadings onto secondary factors (e.g., the order facet of Conscientiousness cross-loads onto Neuroticism), and some traits (e.g., integrity or ambition; Stanek & Ones, 2018) might best be conceptualized as *compound* traits that reflect blends of multiple Big Five domains.

Cross-Cultural Personality Research

An active area of personality research concerns the cross-cultural generalizability of personality structure, as well as personality development and relations of personality traits with behavior and work and life outcomes. A large body of evidence supports the cross-cultural generalizability of personality research. For example, McCrae and colleagues found that Big Five measures developed in the United States showed similar structures when translated and applied in diverse cultures (McCrae et al., 1998, 2005; McCrae & Costa, 1997; Rolland, 2002). Similarly,

Yamagata et al. (2006) found that the genetic structure of a Big Five measure was consistent across cultures, and a series of large-scale studies and meta-analyses have found cross-cultural consistency in patterns of correlations between the Big Five traits and many important outcomes, such as job performance (Barrick et al., 2001; Oh, 2009; Salgado, 1997; van Aarde et al., 2017) and risky sexual behavior (Schmitt, 2004). Taken together, these findings suggest that personality structure and many aspects of personality–behavior relations might be regarded as cultural universals.

Despite these findings supporting cross-cultural generalizability, there remains some concern as to whether the Hierarchical Big Five Model is sufficient to capture all relevant traits across cultures. For example, several researchers have suggested that a six-dimensional factor solution replicates more consistently across cultures (Ashton et al., 2004; Saucier, 2009; but cf. DeYoung, 2010, 2015, for critiques of over-reliance on the lexical hypothesis as a sole criterion for evaluating personality structure). Conversely, in a series of lexical studies across diverse cultures, De Raad et al. (2010) concluded that only three factors (Extraversion, Agreeableness, Conscientiousness) replicated consistently across Western and non-Western cultures; Neuroticism and Openness replicated less clearly across cultures and in some cases appeared to reflect distinct culture-specific constructs.

Cheung et al. (2011; see also John et al., 2008; Rolland, 2002) characterized this discrepancy between the strong generalizability of questionnaire-based personality structures and weaker generalizability of lexical structures as a distinction between "etic" ("external/objective", culturally-comparative) versus "emic" (indigenous, culturally-informed) approaches to personality measurement. Although structures and nomological networks of measures developed in one cultural context may be consistent when the measure is translated and transported to another context (etic approach), these measures may omit important personality traits that are uniquely salient to specific cultures (emic approach). These authors argue that such omissions are particularly likely for non-

American, non-European cultures (e.g., Asian and sub-Saharan African cultures). Like many areas of psychology, personality research has been concentrated on samples from so-called "WEIRD" (Western, Educated, Industrialized, Rich, and Democratic) societies (Arnett, 2008; Henrich et al., 2010).[2] These societies tend to be more individualistic than other cultures, and personality structures identified in these contexts may miss traits related to social relationships that are more relevant in more collectivistic societies. Cheung et al. called for studies to combine etic and emic approaches to identify personality structures that are comparable across cultures but that also include potential culture-specific traits (e.g., Burtăverde et al., 2018; Ion et al., 2016).

Proponents of the combined emic–etic approach to personality assessment argue that measures that include culture-specific traits provide incremental validity above and beyond the Big Five taxonomy (e.g., Burtăverde et al., 2018; Xie et al., 2016). By integrating universal and culturally variable dimensions of personality, these researchers propose that a richer concept of the human personality structure can be derived. Indigenous personality scales are also thought to decrease or eliminate cross-cultural differences in personality scores, while increasing the validity of personality measurement by filling the "blind spot" found in Western trait measures (F. M. Cheung, Cheung, Wada, et al., 2003). A prominent example of combined emic–etic research is the development of the Cross-Cultural (Chinese) Personality Assessment Inventory (CPAI; F. M. Cheung et al., 1996).

The Cross-Cultural (Chinese) Personality Assessment Inventory (CPAI)

Development and structure. The developers of the CPAI sought to use a combined emic–etic approach to construct a set of personality constructs that assess both indigenous personality traits that are informed by Chinese culture and traits derived from scales developed in the West (Cheung et

[2] De Fruyt and Wille (2013) also noted that most *etic* personality research examining personality trait relations with work outcomes has also focused on Western, industrialized, and economically highly-developed societies and called for more attention to applied personality research in diverse cultural and economic contexts.

al., 1996). The researchers developed an initial pool of 150 personality descriptors from a variety of sources, including focus group folk descriptions of personality, contemporary Chinese literature, relevant Chinese proverbs, and the psychological literature. These descriptors included a variety of traits recognizable from existing personality research (e.g., sociability, dependability), as well as traits related to interpersonal interactions that the authors argued were not well-reflected by existing measures. From this initial pool, the authors wrote items to construct 26 normal personality scales, with 28 in the revised CPAI-2, (F. M. Cheung, Shu Fai Cheung, et al., 2008), 12 clinical scales, and 3 validity scales. The 28 normal personality scales load onto four broad factors: Social Potency/ Expansiveness, Dependability, Accommodation, and Interpersonal Relatedness (descriptions of each scale organized by its respective factor are shown in Table 1). The first three factors resemble the three culture-general dimensions identified by De Raad et al. (2010; Extraversion, Conscientiousness, Agreeableness), whereas the fourth, Interpersonal Relatedness, is argued to represent an emic personality factor constructed of indigenous culturally-relevant scales, such as *Harmony* and *Renqing (Relationship Orientation)*. Notably, some indigenously-derived scales load onto the three culture-general "etic" factors (e.g., the *Face* scale loads onto Dependability).

In addition to Traditional and Simplified Chinese, the CPAI-2 has been translated into English, Korean, Dutch, Romanian, and Vietnamese (F. M. Cheung et al., 2011), and has been shown to hold its four-factor structure in cross-cultural samples, especially among Asians and Asian-American participants, but also including European-American samples (F. M. Cheung, Cheung, Leung, et al., 2003; S. F. Cheung et al., 2006; E. J.-L. Lin & Church, 2004). Based on these findings, the CPAI-2 was renamed from the Chinese Personality Assessment Inventory to the *Cross-Cultural* Personality Assessment Inventory to reflect the cross-cultural relevance of some of its indigenously derived trait scales (F. M. Cheung, Shu Fai Cheung, et al., 2008).

Convergence with the Big Five. As noted above, three of the higher-order factors in the CPAI resemble Extraversion, Agreeableness, and Conscientiousness in the Big Five taxonomy. In joint factor analyses of scales from the CPAI and the NEO PI-R or NEO FFI (Costa & McCrae, 1992), Cheung et al. (2001) found that scales from CPAI Interpersonal Relatedness (IR) factor tended to load onto separately from NEO scales (though notably the IR scales showed strong cross-loadings onto Agreeableness). Based on these results, Cheung et al. concluded that the IR factor assesses emic personality features related to instrumental interpersonal relations that are unique to Chinese or collectivist cultures.

In their joint factor analyses, Cheung et al. (2001) found that the NEO Openness scales did not load with any of the CPAI scales. Following these results, the authors added additional Openness-related scales during the revision of the CPAI to create the CPAI-2 (F. M. Cheung, 2006), but subsequent factor analyses found that these scales did not load onto a distinct Openness factor. Instead, most Openness scales loaded onto the Extraversion (Social Potency/Expansiveness) factor (cf. De Raad et al., 2010). Scales assessing Openness facets with cross-loadings onto Agreeableness (e.g., tolerance, emotions [called Interpersonal Sensitivity on the CPAI-2]) loaded most strongly with the Agreeableness (Accommodation) or IR CPAI factors (F. M. Cheung et al., 2011). Based on these results, the authors concluded that the Openness factor was more relevant for Western populations than for the Chinese culture. In another study to explore the relevance of the Openness dimension within Chinese contexts, indigenously-derived Openness items were added to the CPAI-2, leading to the expansion of the Social Potency/Expansiveness, Accommodation and IR factors when it was found that some of the Openness scales load onto them (F. M. Cheung, Shu Fai Cheung, et al., 2008).

Together, these findings have led the developers of the CPAI to conclude that it measures a unique culturally-relevant personality structure that is particularly suited to non-Western, more

collectivistic contexts. However, in a joint factor analysis in a South African sample conducted by a researcher other than the scale developers, the CPAI has shown factor structures more consistent with traditional Big Five structures, with Openness-related CPAI scales loading with other Openness scales and IR scales loading with Agreeableness, Neuroticism, or Conscientiousness scales, rather than forming a separate IR factor (Laher, 2015). Furthermore, studies examining the convergence of the CPAI-2 with the Big Five taxonomy have heretofore only assessed the Big Five taxonomy using the NEO PI-R or NEO FFI. Although these inventories are widely used, their coverage of the construct space for some of the Big Five dimensions is limited (especially Openness; the NEO has limited coverage of the Intellect aspect of Openness and related facets DeYoung et al., 2007; Woo et al., 2014). Additional studies using a wider selection of measures emphasizing different aspects of the Big Five domains are needed before strong conclusions about the potential placement of IR-related emic traits in (or apart from) the Hierarchical Big Five Model trait structure can be drawn.

Validity evidence. Several studies have provided evidence to support the validity of the CPAI-2 normal personality scales for predicting important criteria. Ng et al. (2012) showed that the CPAI-2 differentiated Hong Kong university students across study major and career choice groups and suggested the measure could be useful in career counseling. Several studies have also shown correlations of CPAI scales with well-being indicators, such as life satisfaction and psychological problems (Chen et al., 2006; F. M. Cheung, Kwong, & Zhang, 2003).

In terms of workplace criteria, Kwong and Cheung (2003) reported correlations between the CPAI scales and aspects of contextual job performance (e.g., helping, leadership, initiative, persistence) similar to Western measures of Agreeableness, Extraversion, and Conscientiousness (Chiaburu et al., 2011). Similarly, several studies have found that subscales of the *Dependability* factor predict overall work performance similar to other measures of Conscientiousness (S. F.

Cheung et al., 2007; Ion et al., 2016; Liao, 2005; cf. Connelly et al., 2018). Notably, these studies have observed inconsistent relationships with other CPAI-2 scales, including that the emic trait scales (e.g., *Renqing*, *Harmony*), and found that the emic scales did not show incremental validity over the Conscientiousness-related scales. Tyler and Newcombe (2006) also found that the emic CPAI-2 scales did not substantially predict job performance or show incremental validity over a traditional Big Five-like inventory. Lin (2004) reported workplace creative was positive related to the emic *Renqing* scale and negatively related to the *Veraciousness* scale. In general, the range of workplace criteria used to support workplace applications of the CPAI-2 has been limited, especially for specific performance dimensions such as technical, contextual, and counterproductive performance (Campbell & Wiernik, 2015). Before the CPAI-2 can be justifiably used in organizational decision-making, additional evidence showing the CPAI-2 can predict relevant outcomes is needed.

Methodological limitations. The CPAI-2 is increasingly applied in cross-cultural personality research and practice. However, the evidence cited above shows that some of the claims for its validity, such as the distinction of emic traits from the Hierarchical Big Five Model structure or the incremental validity of emic traits over existing constructs, may be questionable and in need of further scrutiny. In addition, the CPAI-2 also has several methodological limitations that have largely gone unaddressed and that limit its utility for many purposes. First, the inventory is long; the CPAI-2 includes 341 items that make up its 28 normal scales and 3 validity scales. A long test will require more time to complete and can decrease test-taker motivation and increase participant fatigue (Ackerman & Kanfer, 2009). Although in general longer scales are more reliable (Francis & Pearson, 1988), overly long batteries can increase rates of inattentive responding or measurement errors (Curran, 2016; Meade & Craig, 2012). The length of the CPAI-2 particularly makes it difficult to combine with measures of other constructs, limiting the range of research questions it can be used to

address and the breadth constructs that can be assessed in counseling, organizational decision-making, clinical evaluation, or other applications. Many Big Five scales are comparatively shorter without reductions in reliability or validity (e.g., the NEO PI-R has 240 items, Costa & McCrae, 1992; the BFI-2 has 60 items, Soto & John, 2017).

Second, studies using the CPAI-2 have reported low reliability values for some of the scales. For example, Ng et al. (2012) reported coefficient α as low as .47 for the *Renqing* scale, with a median α = .74. Cheung et al. (2008) reported α ranging.43–.84 (median α = .68). Lin and Church (2004) reported an α as low as .22 for one of the scales in their European sample (median α = .59). These low reliabilities raise questions about the appropriateness of using the CPAI scales for applied decision-making. They also potentially call into question the validity of previous conclusions about the CPAI structure. For example, it may be that the IR scales load separately from the Big Five traits due to excessive item heterogeneity, rather than because IR reflects a distinct personality dimension.

Limitations of proprietary scales. A final set of challenges with using the CPA-2 stems from its proprietary nature. Historically, most personality scales were proprietary and commercial. This slowed development of personality assessment research in several ways (Condon & Revelle, 2014; Goldberg, 1999). First, the cost of commercial inventories could be prohibitive for many researchers. Second, many test publishers closely hold scoring keys and sometimes even item content, precluding independent researchers from conducting item-level analyses or validity studies, as well as from using subsets of the full inventory. Third, publishers often have incentives to maintain a measures' status quo; there is frequent interest in keeping test structure, content, and scoring procedures static, limiting refinement of scales and precluding incorporation of advances in personality science into existing measures. Since the advent of public domain and permissively-licensed inventories, such as measures based on the International Personality Item Pool (Goldberg et

al., 2006), personality research has exploded, leading to rapid advances in understanding of the complexity of personality trait structure (Condon et al., 2017; DeYoung et al., 2007), as well as enabling many more researchers to assess personality constructs in their studies. The proprietary nature of the CPAI-2 limits opportunities for other researchers to investigate the structure and validity of the emic personality traits it assesses, particularly alongside other trait measures or in combination with assessments of important work and life outcomes. Public domain measures are needed to open assessment of these emic traits and allow their deeper integration into broader cross-cultural personality science.

The Current Study

In this study, we address several of the limitations discussed above to create a shorter, public-domain version of the CPAI-2. Short public domain scales will permit investigation of the CPAI-2 constructs in a wider range of contexts and allow researchers to probe the factors deeply, especially as part of larger surveys. A shorter scale will enable researchers to use the CPAI-2 to test richer theoretical hypotheses. In particular, a shorter inventory will enable wider tests of the critical hypotheses that the CPAI-2 emic scales assess culturally unique constructs. Much of the existing evidence related to this hypothesis has been presented by the CPAI developers and in only a limited range of cultural contexts; shorter and public-domain measures of these traits will enable more researchers to critically examine the cultural status of these constructs.

For practical applications in counseling or organizations, a shorter measure will enable the CPAI-2 to be included along with measures of other relevant competencies and characteristics to better guide decision-making. Moreover, the low reliabilities of many CPAI-2 scales, despite their length, suggests that these scales are heterogeneous in content. By developing short, more focused scales, with items that more tightly capture construct definitions, it may be possible for scales to not only be shorter, but also more reliable and coherent (cf. Saucier, 1994). We will use multiple analytic

methods to develop short-form scales and compare the reliability and validity of scales developed using different approaches.

In addition to developing public domain scales, we will also gather additional evidence to test the construct and criterion-related validity of the CPAI-2 constructs, as measured by both the original and public domain measures. We will conduct factor analyses to evaluate whether our new scales fit the CPAI factor structure. We will also examine convergence of the original and public domain CPAI-2 scales with a traditional measure of Big Five domains and facet traits that has not previously been employed in CPAI-2 validation and which is designed to balance coverage of the major lower-order components of the Big Five domains that have been identified in recent personality research (Soto & John, 2017). We will also examine relations of the original and public domain CPAI-2 scales with a range of important outcomes, including contextual and counterproductive job performance (Campbell & Wiernik, 2015) and psychological distress (cf. Berkel & Kathryn, 2009; Klein et al., 2011).

Challenges in developing short-form measures. There are several important challenges that must be considered when developing short-form scales. These challenges have been discussed by previous developers of short personality scales (Gosling et al., 2003; Saucier, 1994; Soto & John, 2017), as well as in critiques of ultra-short personality scales (Credé et al., 2012; Hofmans et al., 2008). First, compared with longer personality scales, short scales can have lower reliability, particularly as indexed using coefficient α or other internal consistency estimators. Second, short form scales may suffer a loss of fidelity, with important trait content being omitted and changing the meaning of scores (Credé et al., 2012; Hofmans et al., 2008; Saucier, 1994). Third, removing items can change the correlations among scales in an inventory. This can be positive (e.g., if artefactual

intercorrelations are reduced; Saucier, 1994), but it may also increase the impact of response biases and method effects, inflating correlations and harming scales' factor structure.

To address these challenges, developers of short-form measures must ensure that reduced scales retain adequate coverage of constructs' full scopes, show good convergence with longer scales, preserve meaningful factor structures. Regarding reliability, it is important to evaluate short-form measures using appropriate reliability estimators. Internal consistency estimators, such as coefficient α, are highly influenced by test length and can in many ways be regarded as indices of item redundancy; a measure will show high internal consistency if it has many items providing overlapping information (high intercorrelations). Short-form scales are often designed to remove redundant items, so internal consistency may not provide an accurate representation of the reliability of the scale. Instead, estimates of the coefficient of equivalence (e.g., parallel forms reliability) and coefficient of stability (e.g., test–retest reliability) may provide more realistic indices of the amount and consequences of measurement error in a measure (Wiernik & Dahlke, 2020). Yarkoni (2010) demonstrated that parallel forms and test-retest reliability coefficients can be very highly for short measures, even if internal consistency estimates are comparatively low.

For the current study, we will evaluate our public domain short-form analogues of the CPAI-2 scales in terms of their factor structure, convergence with the original scales, and relationships with criterion variables. The broad aim of this study is to provide guidance on public domain assessment of the CPAI-2 constructs. We will be using two different approaches for the two different sets of CPAI-2 scales: etic and emic. The etic constructs assessed by the CPAI are well-captured by existing public domain scales. However, for the emic scales, there are no clear existing analogues. As such, we will be using a nomological web clustering approach (Hough & Ones, 2001; Stanek & Ones, 2018) to identify public domain analogues to the etic scales, and we will also develop new public

domain items to assess the emic scales. Our study includes two phases. Phase 1 involves the analogue scales construction process, and Phase 2 involves the new emic scales validation process.

Chapter 2: Method

Phase 1: Analogue Scales Construction Process

The CPAI-2 includes two broad types of normal personality content scales—etic scales, which are measures of traits commonly assessed by Western-developed personality inventories (e.g., Extraversion versus Introversion), and emic scales, which assess personality trait constructs that are purportedly unique the Chinese or other non-Western contexts (e.g., Face, Renqing). We adopt different strategies to develop public domain analogues for these two categories of scales.

Etic scales. The etic scales are explicitly similar to existing measures of well-known, understood, and documented personality constructs. The CPAI etic scales are: Novelty, Diversity, Divergent Thinking, Leadership, Aesthetics, Extraversion vs. Introversion, Enterprise, Responsibility, Emotionality, Inferiority vs. Self-Acceptance, Optimism vs. Pessimism, Meticulousness, Internal vs. External Locus of Control, Veraciousness vs. Slickness, Interpersonal Tolerance, Logical vs. Affective Orientation, and Practical Mindedness. Each of these scales assesses a construct that is also captured by a range of existing personality scales. For example, the Extraversion vs. Introversion scale assesses the construct of *sociability*, a trait also assessed by the NEO PI-R Gregariousness scale and the Hogan Personality Inventory Sociability scale, among others (Stanek & Ones, 2018).

Accordingly, for the CPAI-2 etic scales, we adopted the nomological web-clustering approach (Hough & Ones, 2001; Stanek & Ones, 2018) to identify existing public domain measures that align with the constructs assessed by the CPAI-2 scales. In this approach, we examined scale descriptions, example item content, and patterns of correlations with other measures reported in

previous studies (e.g., F. M. Cheung et al., 2001; F. M. Cheung, Shu Fai Cheung, et al., 2008; Ng et al., 2012; summarized in Table 2) to determine which personality construct is assessed by each CPAI-2 etic scale. We then consulted the taxonomy described by Stanek and Ones (2018), who applied this approach to a large number of personality scales, to identify existing public domain scales capturing the same constructs. There are many existing scales that could be used to measure these constructs, and choice among them is inherently a subjective process. Before selecting scales, we set out a priori criteria to guide our choices. We selected scales based on (1) their semantic similarity to the CPAI-2 scales, (2) their reported reliability, (3) their length (preferring shorter scales), and (4) their frequency of application (i.e., widely-used IPIP items scales were preferred over less popular scales). A list of CPAI-2 etic scales matched with suggested analogue public domain scales is provided in Table 3.

We adopted the nomological web-clustering approach for etic scales for two reasons. First, the approach is well-established as a method for identifying commensurability of traits assessed by different inventories (e.g., in grouping personality measures in meta-analyses; Barrick & Mount, 1991; Dudley et al., 2006; Salgado, 2003). Second, due the proprietary nature of the CPAI-2, we are not able to administer subsets of the CPAI-2 items. This limits the number of additional scales we can administer to participants alongside the 341-item CPAI-2. Accordingly, we chose to prioritize development of new public domain measures of the CPAI-2 emic traits and assessment of the construct and criterion-related validity of the public domain and existing CPAI-2 scales.

Although we have identified the analogues for the etic scales here, we will not be administering any of these scales during the scale validation process to limit survey fatigue. If we were to administer all items from all analogue scales along with the entire CPAI, the test would be incredibly long and hence, difficult to administer. As such, we have identified and select the most

appropriate public domain analogues to the CPAI-2 etic scales and have listed them in Table 3 to provide guidance for future researchers.

Emic scales. The emic scales were constructed to assess purported culture-specific traits, so there are no clear analogues for these scales among existing public domain personality measures. Accordingly, we propose to develop new public domain items to assess the CPAI-2 emic traits. We will follow the recommendations provided by Boateng et al. (2018), who describe scale development and validation as a three stage process composed of nine steps.

In the first stage, we identified the target domain and generated items. In our case, the domains of interest are the 11 CPAI emic traits: Face, Family Orientation, Defensiveness, Graciousness vs. Meanness, Self vs. Social Orientation, Traditionalism vs. Modernity, Renqing (Relationship Orientation), Social Sensitivity, Discipline, Harmony, and Thrift vs. Extravagance. We generated candidate items to assess each construct based on scale descriptions and dimensions derived from F. M. Cheung et al. (2008) and S. F. Cheung et al. (2013).

We used an item writing worksheet (Appendix C) to train research assistants on item generation and to standardize the item generation process for all scales. Research assistants with prior experience in personality assessment and familiarity with Southern, South-eastern, and Eastern Asian cultures generated English-language items for each scale using the worksheet. These are some of the cultures for which the CPAI developers have suggested the CPAI traits are relevant. After candidate item pools were developed, the first and last authors of this paper were responsible for reviewing the items for redundancy, content relevance, and clarity, and for selecting the final pool of items for validation. These finalized scales were then piloted by administering all items to several research assistants. Research assistants who were not involved in item writing read each item reviewing them for clarity, readability, and flow. Items that research assistants were unsure how to respond to or

those that were found to be unclear were reported. Items found to be unclear by the first and last authors and by research assistants were revised, and pending issues were resolved by meeting.

Then, authors of the current paper who have familiarity with personality measurement research and are fluent in both Chinese and English translated the items into Chinese. We used back-translation to ensure linguistic equivalence of the English and Chinese items (ITC, 2005) and conflicts of opinion were resolved through discussion. After translation, a Chinese-speaking research assistant not involved in the translation process took the Chinese version of the survey and identified areas lacking clarity. These were then revised accordingly.

Phase 2: New Emic Scales Validation Process

Participants. Two samples were collected: one from the U.S and one from China. The U.S. sample comprised of a university student sample recruited from the authors' institution and a sample of working adults recruited using Amazon Mechanical Turk (MTurk). MTurk data was collected using CloudResearch (formerly called TurkPrime; Litman et al., 2017), a crowdsourced data collection platform that allows for multiple data quality screening techniques. MTurk participants were compensated with $5 for every 30 minutes spent on answering items, with a total of $15 for completing the study. The selection criteria for U.S. samples was that all participants must be aged 18 or older, speak English as a first language, live in the U.S., and be currently employed or have been employed in the past.

The Chinese sample was recruited using Credamo, an online survey recruitment platform based in China that is similar to Profilic in the U.S. Credamo participants were compensated with 60 Chinese Yuan to complete the full set of scales. Selection criteria for the Chinese sample were similar to those outlined above with the exceptions being that participants must speak Mandarin as a first language and must live in China.

To facilitate examination of CPAI-2 relations with work behaviors, all participants were additionally required to either be currently employed or to have been employed in the past. Respondents who were identified as exhibiting careless/inattentive responding (2016) were excluded from the analysis.

Sample Size Planning. Each of the two analytic approaches employed in this study to create short public domain CPAI-2 analogue scales rely on correlations among items. As a result, the fundamental sample size consideration for these analyses is the sample size at which the magnitude of sample correlation coefficients stabilizes. Schönbrodt and Perugini (2013) conducted a series of Monte Carlo simulations to identify the *point of stability* for population correlations of different magnitudes. The *point of stability* is defined as the sample size where sample deviations from the true population fall within a specified small range (e.g., ±.10). Based on the coefficient α values for the CPAI-2 scales reported by Ng et al. (2012), the average mean interitem correlation for items within each CPAI-2 scale is $\bar{r}_{ij}$ = .21 (median = .22, *SD* = .08). These values are somewhat smaller than correlations typically observed among personality items (e.g., $\bar{r}$ = .48 among items from the same narrow trait scale; $\bar{r}$ = .32 among items from the same broad Big Five trait scale; $\bar{r}$ = .12 among self-report personality items overall; D. M. Condon, personal communication, 22 January 2020, based on data from Condon et al., 2017). Accordingly, based on Schönbrodt and Perugini's results for ρ = .20, we can expect stable interitem correlation estimates with ±.10 (80% confidence) at *N* = 238. Ng et al. (2012) reported α < .65 for 5 scales (*Practical Mindedness*, *Face*, *Renqing*, *Discipline*, *Thrift vs. Extravagence*; average $\bar{r}_{ij}$ = .10 for these scales). Based on Schönbrodt and Perugini's results for ρ = .10, we can expect stable interitem correlation estimates with ±.10 (80% confidence) at *N* = 252 for these less internally-consistent scales.

Previous power estimates for factor analysis methods have not provided definitive

recommendations for sample size requirements (Orlando, 2004). However, there are general guidelines based on simulation evidence that can be used to determine an appropriate sample size. For IFA analyses, Cappelleri et al. (2014) suggest that a sample size of ≥ 500 is ideal for 2-parameter logistic (2PL) models, but a smaller sample could still provide useful information for model estimation. Additionally, other research has suggested that smaller samples of as low as $N = 200$ can provide accurate parameter estimates (e.g., Orlando & Marshall, 2002; Thissen & Steinberg, 1986). Furthermore, Şahin and Anıl (2016) demonstrated that accurate *a* and *b* parameter estimates can be obtained when $N = 250$. This is also supported by Goldman and Raju (1986) who suggest that a sample size of ≥ 250 is sufficient for IFA 2PL model estimation. Notably, these analyses generally assume that item communalities are high. When factors are well-determined with highly discriminating items, sample size requirements for linear factor analysis and item factor analysis are lower than when item discrimination parameters are lower (Cappelleri et al., 2014; MacCallum et al., 1999; Tabachnick & Fidell, 2018).

Previous applications of genetic algorithms to abbreviate personality measures have found good performance with similar sample sizes (e.g., *N*s = 114 [cross-validated using retest data]; 857 [divided in half for cross-validation]; 1,590 [divided in half for cross-validation]) (Eisenbarth et al., 2015; Yarkoni, 2010). However, there are no robust simulation studies using genetic algorithms to test for required sample sizes, so we conducted an iterative simulation analysis to test a range of sample sizes. We used the genetic algorithm to shorten a 300 item IPIP version of the NEO-PI-R (Goldberg, 1999) using data from Johnson (2014) at a range of sample sizes from 100 to 1000 to determine the appropriate sample size for the current study. The data consists of 307,313 cases of responses to the IPIP-NEO-300. The IPIP-NEO-300 is a public domain 300-item measure of the Big Five personality traits arranged in 30 facets similar to the NEO-PI-R. Since the data from this study is

publicly available and the length of the scale is similar to that of the CPAI-2, we used the data to test different sample sizes. We repeatedly drew sample sizes of 100, 250, 375, 500, 750 and 1000 at random from the large dataset and simulated the GA 100 times for each sample size. From the output, we extracted the convergent correlations for each of the 30 facets and then calculated the average convergent correlations for each sample size. Mean convergent correlations for each sample size tested along with the range and 95% confidence interval are presented in Table 4. Results show that the convergent correlations remained consistent at around 0.65 across the different sample sizes tested (with a range of 0.637 – 0.655). The mean convergent correlations for each of the sample sizes are as follows: at N = 100, M = 0.655, SD = 0.066, 95% CI = [0.642, 0.668]; at N = 250, M= 0.650, SD = 0.052, 95% CI = [0.637, 0.663]; at N = 375, M = 0.645, SD = 0.046, 95% CI = [0.633, 0.658]; at N = 500, M = 0.642, SD = 0.043, 95% CI = [0.629, 0.655]; at N = 750, M = 0.638, SD = 0.042, 95% CI = [0.626, 0.651]; at N = 1000, M = 0.637, SD = 0.041, 95% CI = [0.624, 0.650]. The overlapping confidence intervals of the means suggest that there is no significant difference in convergent correlations across the different sample sizes used in the GA. The GA functions as well with small samples as with large samples.

Based on this empirical test, Schönbrodt and Perugini's simulations and guidelines in the literature, we propose to collect data from 400 participants in each of the U.S. and Chinese samples. This sample size reflects the final sample size to be included in our analyses.

The final sample included a total of N = 918 participants with 313 participants from China and 605 participants from the US. Full sample demographics are presented in Table 5.

Measures. U.S. participants completed all measures in English. Chinese participants completed all measures in Simplified Chinese.

CPAI-2. The Cross-Cultural (Chinese) Personality Assessment Inventory-2 (F. M. Cheung,

Shu Fai Cheung, et al., 2008) is a measure of personality containing 341 items organized into 28 normal (non-clinical) personality scales. Respondents were asked to rate the extent to which they agree or disagree with each statement describing a behavior characteristic of personality traits on a 5-point Likert-type scale from 1 (least) to 5 (most). The clinical scales of this measure were not used.

BFI-2. The Big Five Inventory-2 (BFI-2; Soto & John, 2017) was used to measure the Big Five domains and their major lower-order traits. The BFI-2 is a widely used 60-item personality measure which was developed to provide brief but reliable assessment of both the Big Five traits and their major lower-order components (15 narrow traits). The facets assessed by the BFI-2 were chosen based on narrow traits that have been empirically identified in recent decades of personality research (cf. the rationally-derived scales of the NEO PI-R). The BFI-2 items are written to be understandable to readers with a range of English reading ability levels. The BFI-2 has shown to have high reliability and validity, and it has been translated and validated in a variety of languages. Respondents were asked to rate the extent to which they agree or disagree with each statement describing a behavior characteristic of personality traits on a 5-point Likert-type scale from 1 (least) to 5 (most). Reliability of the scales in this study ranged from $\alpha = 0.88$ to 0.93.

Public domain CPAI-2 analogue scales. After generating new items to assess the 11 emic CPAI-2 constructs for which no clear existing public domain measure can be identified, 77 items were retained to be administered as part of the total battery of items.

Criterion-related validity. Criterion-related validity of the original and public domain forms of the CPAI-2 was assessed using measures of criterion constructs important to applied organizational and clinical assessment, including contextual job performance/organizational citizenship behaviors (Borman & Motowidlo, 1997; Chiaburu et al., 2011), counterproductive work behaviors (Ones & Dilchert, 2013), and general psychological well-being/distress (Henry &

Crawford, 2005). These criteria are not only important for their respective fields, but have also been frequently identified as key outcomes associated with personality traits (Connelly et al., 2018; Ozer & Benet-Martínez, 2006).

Organizational citizenship behavior. Organizational citizenship behavior (OCB) was assessed using the 20-item OCB Checklist (OCB-C) (Fox et al., 2012). The scale assesses interpersonal OCB (helping coworkers, cooperation, etc.) and organizationally directed OCB (supporting organizational goals, representing the organization well). The OCB-C is a widely used measure of OCB that has been shown to have good reliability and validity and has good convergence with other OCB measures. In this study, the OCB-C had good reliability ($\alpha = 0.93$).

Counterproductive work behavior. Counterproductive work behavior (CWB) was measured using the 45-item CWB Checklist (CWB-C) (Spector et al., 2006). The CWB-C is a widely-used CWB measure that assesses 5 types of counterproductive behaviors – Abuse, Production Deviance, Sabotage, Theft and Withdrawal. The CWB-C has been shown to have good reliability and convergence with other CWB measures. In this study, the CWB-C had good reliability ($\alpha = 0.96$). [3]

Psychological distress. The Depression, Anxiety, and Stress Scale (DASS-21) (Henry & Crawford, 2005) measures participants' experiences of psychological stress, anxiety, and depression. Respondents reported the frequency and severity of 21 negative emotional symptoms. The DASS-21 has 3 subscales measuring depression, anxiety, and stress, each with 7 items. The DASS-21 is a widely-used measure of general psychological distress. It demonstrates good psychometric properties and has an established set of norms that can provide a useful point of comparison for our current sample (Henry & Crawford, 2005). The DASS-21 appears to work well not just as a clinical instrument but also for assessing normal-range levels of psychological distress. In this study, the

[3] Due to a clerical error, the CWB scales were not administered to a portion of a US participants and thus we had to oversample from the US in order to collect full additional response sets including the CWB scales.

DASS-21 had good reliability overall (α = 0.97) and across the three depression, anxiety, and stress scales (α = 0.96; 0.91; 0.91).

Data Collection. This study employed self-report measures of personality, job performance, and psychological health. All measures were completed online. Participants were recruited through the student study participation pool, MTurk, or Credamo platforms. After electing to participate, participants completed the informed consent process, followed by the study measures. All participants completed the CPAI-2 and all newly-generated items. In addition, half of the participants in each sample were randomly assigned to complete the BFI-2, whereas the other half completed the criterion measures—OCB-C, CWB-C, and DASS-21. Scale order and item order within each scale were randomized to reduce potential order effects.

Before data analysis, respondents were screened for inattentive responding using the multi-pronged approach recommend by Curran (2016). This approach considers a variety of potential careless responding, including self-report questions asking participants whether they responded carefully and accurately; examination of response times; long-string analysis (repeatedly answering with the same option for many items); and person–total correlations (correlation between a person's response vector and the vector of sample average responses for each item). In addition, we identified potential influential cases by calculating generalized Cook's distance (Cook, 1986) for a unidimensional factor analysis of each included scale. For the self-report attentiveness questionnaire, we regarded individuals who answer "No" to the question, "In your honest opinion, should we use your data in our analyses in this study?" as inattentive responders and removed them from analyses. For response times, we removed individuals with implausibly fast completion (defined as more than 2 MAD [median absolute deviation from the median] less than the median response time) of each scale and overall (expected total completion time is 50–90 minutes) from analyses. For long-string

analysis and person–total correlations, we regarded individuals with responses more than 2 standard deviations from the sample mean as inattentively responding and removed them from analyses. For Cook's distance, we regarded individuals with $D \geq 2.0$ for a scale as potential influential cases and examined results using the scale with and without these influential cases. Detection of inattentive responding and influential outliers is inherently a subjective process, so we conducted all analyses using both the full and screened datasets and did not note any discrepant results. The results shared in this manuscript only include participants that were not excluded for careless/inattentive responding.

Data Analysis

The goal of this study is to develop a public-domain short-form of the CPAI-2 that maintains factor structure of and good convergence with the long-form measure, maintains adequate reliability, and maintains strong validity. Our analyses of the new public domain scales proceeded in three stages. First, we selected final items for each new public domain scale using two different item-selection approaches and compared the reliability and structure of measures constructed using each approach. In this stage, we also examined the convergence of our new scales with the original scales, as well as the similarity of the patterns of intercorrelations among our new scales and the original CPAI-2 scales. Second, we examined the convergence of the patterns of correlations of our new scales and the original CPAI-2 scales with Big Five domains and facets as assessed by the BFI-2. Third, we examined the criterion-related validity of the original and public domain CPAI-2 scales for OCB, CWB, and psychological well-being, as well as their incremental validity over the BFI-2 scales.

Evaluating Items for the Public-Domain Scales

We will develop short public domain forms of the CPAI-2 constructs using two approaches. The first, item factor analysis (IFA), is a traditional technique for scale construction and involves rationally selecting the best performing items based on their performance as indicators of latent

common factor. The second approach is a relatively novel automated method of item selection that uses a machine-learning technique called a Genetic Algorithm to automatically select a best-performing reduced item set according to preset criteria. We describe each of these approaches below.

Item factor analysis. The first analytic approach we will use is item factor analysis (IFA; i.e., nonlinear factor analysis or item response theory; Pritikin et al., 2014). In IFA, item responses are treated as indicators of underlying latent variables, and items are chosen based on their ability to discriminate individuals with higher versus lower levels of the modeled latent variables. Various forms of factor analysis are the most commonly used approach in developing or shortening psychological measures (e.g., Porter et al., 2016; Saucier, 1994). During the development of the CPAI-2, Cheung et al. (1996) used a linear factor analysis model, where item responses are treated as continuous and the item intercept (cf. base rate or difficulty) is assumed to be equal across items. However, a potential limitation of the linear factor analysis model is that it does not consider item difficulties/locations when estimating their relationship with proposed latent factors. An item representing an extreme level of a trait may show a weak correlation with other items (and thus weak linear factor loadings) because of relatively low variance, rather than truly weak relationships with the latent factor. By ignoring item location, linear factor analysis tends to retain only items located near the trait mean (where variance is highest) and remove more extreme items, even though these items may be valuable for assessing respondents with very high or low trait levels. IFA models explicitly model both an item's location and its latent trait discrimination, providing more accurate assessment of the item's relationship with the latent trait. A second potential limitation of linear factor analysis is that it assumes that the points on the Likert-type response scale of the included items are evenly spaced; item factor analysis models such as the graded response model do not make

this assumption and instead estimate a location threshold parameter for each scale point.

Compared to linear factor analysis, in practice, considering item or scale point locations leads to an important difference in how items are selected. Whereas item selection in linear factor analysis typically considers each item separately by selecting the items with the highest factor loadings, IFA item selection focuses on selecting a *set* of items that together provide high information (precise measurement) across a range of trait levels. Thus, not only item discrimination, but also the range of scale point locations, are considered. Selection based on only the factor loading tends to select items located near the center of the trait distribution. This can lead to a peaked test information function, indicating that the test provides reliable measurement near the mean of the trait distribution, but relatively poor measurement at even somewhat extreme trait levels.

To develop public-domain short-form CPAI-2 scales using IFA, for each of the 11 new analogue emic scales, we will fit a unidimensional graded response model (Samejima, 1997) using the *mirt* packge (Chalmers, 2012) in *R* (R Core Team, 2020) to the candidate items. We will evaluate items using their discrimination parameters, item information curves, and item–model fit. We will retain a set of 4–6 items that together provide a high level of total test information (e.g., $I(\theta) \geq 2.0$) across a wide range of trait levels (e.g., $\theta = [-2.0, +2.0]$; cf. Wiernik et al., 2019). In addition to these statistical criteria, we will also consider 3 rational criteria. First, retained items should load onto the common factor in the theoretically expected direction. Second, the final item set should broadly cover the content domain targeted by the scale (i.e., item content should not be overly redundant). Third, the final item set should represent both the high and low poles of the trait, to the extent that the original item set includes both positively and negatively keyed items.[4]

[4] A disadvantage of discrete response IFA models is that estimating multiple item parameters, including discrimination and multiple threshold parameters, can require comparatively larger samples than linear factor analysis models. If the sample size is too small, standard errors for item parameters will be large, making it difficult to evaluate item performance.

This combination of statistical and rational criteria can help to avoid the common pitfall of short scales wherein a limited set of highly similar items is selected. For such high-redundancy scales, high item intercorrelations or discrimination parameters will reflect a combination of both reliable variance and shared item-specific (content sampling) error (Wiernik & Dahlke, 2020). As a result, when such scales show high coefficient α or test information, this value more reflects item content overlap rather than the reliability of the scale as a measure of the intended latent construct. By ensuring broad bandwidth of the reduced item set, we can avoid this problem. Similar combination statistical–rational approaches have previously been used successfully to develop short personality measures (Gerlitz & Schupp, 2005; e.g., Gosling et al., 2003; Hahn et al., 2012; Rammstedt & John, 2007).

We chose to use unidimensional models to construct scales individually rather than simultaneously for two reasons. First, given the hierarchical structure of the CPAI-2 and the large number of narrow trait scales, it would be difficult to identify stable factor solutions if all item were analyzed simultaneously and 11 factors were extracted (cf. Costa & McCrae, 1992). Second, the 11 CPAI-2 scales we are creating new items for do not have clean simple-structure loadings onto higher-order factors. These cross-loadings would make it challenging to interpret the meaning of item factor loadings as a basis for item selection. For example, if all of the items for narrow traits loading onto the IR factor were factor analyzed together, the resulting latent factors would reflect (1) shared variance with the higher-order latent factor, (2) unique variance for the narrow latent trait, and (3) variance coming from the other higher-order factors onto which the scales cross-load. Selecting

In this study, although we will prefer a discrete response IFA model if possible, before retaining such a model, we will examine the standard errors for item parameters. In the case that observed standard errors for item parameters are too large due to inadequate power, we will revert to a linear factor analysis model instead. In such a case, the model will be estimated as a unidimensional minimum residuals factor analysis using the `fa()` function in the *psych* package (Revelle, 2019) in *R*. We will retain a set of 4–6 items that show substantial factor loadings ($\lambda \geq .40$) in the expected direction and which meet our rational criteria.

items with strong loadings on one narrow trait factor and weak loadings onto the other narrow trait factors within a higher-order domain may tend to prioritize the cross-loading variance of the items, rather than the variance that is shared by the narrow traits as part of the higher-order domain. This may have the consequence of decreasing theoretically relevant convergent correlations among narrow traits within a domain (and the reliability of the higher-order scales), as well as increasing the correlations among narrow traits across domains (and among the higher-order scores).

Genetic algorithm. The approach described above is labor intensive. Although IFA uses statistical information, final item selection relies on rational judgments about comparison of item parameters and information curves, content coverage, keying balance, and other factors. In addition, the new scales developed using this method do not necessarily maintain strong convergence with the original scales. Yarkoni (2010) proposed an alternative approach to scale abbreviation that uses a machine learning technique called a genetic algorithm (GA) to automatically select a set of items that optimally balances scale brevity and convergence with the full-length scale. This method iteratively creates and tests possible scoring keys to select the best items out of the candidate pool. A particular strength of the genetic algorithm is that it prioritizes item non-redundancy, avoiding problems that can result from selecting items based on factor loadings if careful attention is not given to maintaining content coverage of shortened scales (Eisenbarth et al., 2015; Yarkoni, 2010). The GA approach has previously been applied to shorten personality (Eisenbarth et al., 2015; Sandy et al., 2014; Yarkoni, 2010) and ability (Schroeders et al., 2016) scales while maintaining good psychometric properties.

The GA approach uses an evolutionary metaphor to select the highest-quality items. In an evolutionary sense, the fittest chromosomes from each generation survive to repopulate the pool of chromosomes in successive generations. Similarly, the GA approach iteratively randomly generates

candidate scoring keys, tests the fitness of keys against a cost function, and then generates a new set of candidate scoring keys based on the fittest keys. This cycle is repeated until successive generations no longer produce better (lower cost) scoring keys.

The GA cost function used to test the fitness of candidate scoring keys is designed to optimize two criteria—(1) minimize the number of items retained, and (2) maximize the fidelity (convergent R^2) of the new scales as proxies for the original scales. These criteria are necessarily in conflict (removing items generally reduces convergence with the full-length scale), so optimizing this function leads the GA to balance these two goals of scale abbreviation. The GA cost function is (Yarkoni, 2010, p. 182):

$$Cost = I \times k + \sum_{i=1}^{s} (1 - R_i^2),$$

where I is the fixed item cost, k is the total number of items retained, s is the number of scales in the inventory, and R_i^2 is amount of variance in the i^{th} scale that is accounted for by the scoring key.

The first term of this cost function ($I \times k$) reflects the penalty associated with longer scales. The second term of the cost function $(1 - R_i^2)$ reflects the penalty associated with non-fidelity (original scale variance not captured by the new scoring keys). The relative weight given to these two criteria is controlled by the item cost parameter, I. Large values for I prioritize inventory brevity by making the cost of adding an additional item outweigh the cost of loss in fidelity. Small values for I prioritize scale fidelity, even at the expense of a longer measure. Because of the cost associated with adding items, the GA favors non-redundant items (i.e., items that are weakly intercorrelated). The GA optimizes scoring keys for all of the scales of an inventory simultaneously, and items are permitted to score on multiple scales; previous applications of the GA to abbreviate personality scales has found that it retained the internal structure of the inventory well, even in cases of some degree of item overlap across scales (Yarkoni, 2010).

Several researchers have found success with using the aforementioned cost function for developing short-form scales using the GA, resulting in well-functioning scales (Nikan et al., 2018; Olaru et al., 2015; Raborn et al., 2020; Sandy et al., 2014; Schroeders et al., 2016). Yarkoni (2010) explains that different values of item cost parameter can result in significant differences in the fidelity-to-brevity ratio. Setting a lower item cost parameter will result in higher convergent validity (fidelity) with the original scale but will retain a larger number of items. Setting a larger item cost parameter will more severely shorten the scales but may also reduce convergent validity with the original scales. In an application of the GA to shorten the Psychopathic Personality Inventory–Revised, Eisenbarth et al. (2015) considered *I* cost values ranging from .02 to .08, with the maximum number of items scored per scale set to 3, 5, 7, or 9. These authors found that these combinations yielded short scales with convergent correlations ranging from $r = .81$ to .95. They ultimately chose to retain the item set selected with parameters $I = .02$ and a maximum of 5 items per scale, arguing that this solution subjectively most effectively balanced convergence and brevity (convergent $r = .91$ with 40 items). We will adopt a similar approach by running the GA with a range of *I* cost values and maximum numbers of items per scale set to 4, 5, or 6. We will then select the values that subjectively provides the best balance between brevity and fidelity. This approach will both allow us to probe the impact of the item cost parameter on the resulting scale lengths and fidelities, as well as provide illustrative guidance for future researchers wishing to apply the GA for scale shortening.

Similar to the IFA approach, we will use the GA approach to develop new public domain analogues for 11 CPAI-2 narrow scales, and we will set the maximum number of items per scale to six. We will use code adapted from Yarkoni (2010) and the *genalg* package (Willighagen, 2005) in *R* to conduct the GA analyses. Example scripts for these analyses are provided in the Supplemental Materials.

Evaluating Short-Form Inventories

We will use the two approaches described above to develop two sets of new short public domain CPAI-2 scales. We will evaluate these inventories in terms of their number of retained items, convergence with full-length scales, structural validity, correlations with Big Five traits, and criterion-related validity.

Convergence with full-length scales and structural validity. The most basic evaluation criterion for public domain analogues of scales is their fidelity as proxies for original measures. We will compare the new scales developed using the two approaches in terms of their convergent correlations with their corresponding original scales, with higher correlations indicating higher quality.

In developing short public domain analogue scales, our goal is to preserve the internal structure of the CPAI-2 in terms of correlations among the scales. A potential concern with short-form scales is that the nuanced patterns of scale interrelations may be lost. For example, on a short-form inventory, facet scales of a higher-order factor may lose their discriminant validity and instead primarily reflect only shared variance with the higher-order factor. This may be especially likely if items are scored on multiple scales (as is possible with the GA approach; but cf. Yarkoni, 2010). Conversely, by removing low quality items, short scales may instead improve the clarity of factor structures by reducing cross-loadings of items or correlations among theoretically-distinct constructs (Saucier, 1994).

We will evaluate the structural validity of the new public domain scales by examining the pattern of correlations among the 11 new narrow scales, as well as the similarity of the new scales and original scales in terms of their patterns of correlations with the other CPAI-2 scales. Quality of the new scales will be determined by the degree to which the magnitudes of scale intercorrelations

reproduce the correlations among the corresponding original scales (as well as potentially reduced correlations among theoretically-distinct constructs; Saucier, 1994).

Correlations with the Big Five. We will examine correlations of each new public domain CPAI-2 scale with the BFI-2 domain and facet scales. Quality of the new scales will be determined by the degree to which the new scales reproduce the patterns and magnitudes of Big Five and facet correlations shown by the original scales. These correlations will further permit additional tests of the discriminant validity of the CPAI-2 emic scales, and the IR factor in particular, from the Big Five trait hierarchy using a new measure which has not been applied in previous tests of this cultural specificity hypothesis and which incorporates more recent advances in understanding of lower-order facets of the Big Five (Soto & John, 2017).

Criterion-related validity. To estimate the criterion-related validity of the public domain CPAI-2 scales, we will examine correlations of each new public domain CPAI-2 scale with the measures of OCB, CWB, and psychological well-being. Quality of the public domain measures will be determined by the degree to which the public domain scales reproduce the pattern and magnitude of the criterion correlations displayed by the original scales. These correlations will further expand our understanding of the nomological network of the CPAI-2 constructs by examining concurrent validity with criterion constructs (e.g., dimensions of job performance) that have not been widely considered in existing CPAI-2 validation research.

Reliability. We note that assessment of reliability of short-form measures is difficult. By linking reliability with shared item variance (redundancy), internal consistency estimates such as coefficient α or IRT-based reliability are likely to underestimate scale reliability for short or heterogeneous scales (Osburn, 2000). In particular, the GA scale abbreviation process is explicitly designed to remove redundant variance from a measure to optimize both fidelity and brevity, so a

short-form scale with high convergent validity and low internal consistency is the ideal outcome (Yarkoni, 2010). For short measures, parallel forms or test–retest methods provide more appropriate reliability estimates (Gosling et al., 2003; Wiernik & Dahlke, 2020). Measure reliability can also be inferred by examining patterns of correlations with external variables, such as the Big Five traits or criterion constructs. As reliability is the theoretical ceiling for validity (Nunnally, 1978), if a short-form scale shows similar levels of validity as its long-form counterpart, this implies that the scales have similar levels of reliability. With these caveats in mind, for comparison with previous studies developing short-form scales, we will estimate internal consistency for the short-form CPAI-2 scales, including coefficient α, model-based internal consistency estimates (IRT empirical reliability for IRT-based scales, Chalmers, 2012; for the GA scales, we will fit unidimensional graded response IRT models and estimate empirical reliability). Considering our use of Likert-type response scales, we will additionally compute ordinal forms of coefficient α for all scales.

Chapter 3: Results

Scale means, standard deviations, and Cronbach's α of all measures are presented for the overall sample and the two subsamples in Table 6.

Scale Characteristics of all new CPAI items

The full set of newly created analogue items for assessing the emic CPAI-2 traits will be referenced to as "new CPAI" items and scales. The means, standard deviations, Cronbach's alpha, item-scale correlations, IRT discrimination and intercept parameters, and marginal reliability for all new CPAI items and scales are presented in Table 7. Scale means for all new CPAI scales are presented in Table 6. Additionally, Figure 1 compares the distribution of new CPAI scale scores across the Chinese and US samples.

Convergence with full-length scales. Correlations among the new CPAI scales and the equivalent CPAI-2 emic scales are presented in Table 8 and Figure 2. In addition, Figure 3 illustrates the magnitude of and direction of the correlation between each equivalent scale pair, i.e., the same scales as assessed by the new CPAI items and the CPAI-2 items.

Results show that all new CPAI scales are positively and significantly correlated with their CPAI-2 counterparts with correlations ranging from r [95% CI] = 0.25 [0.19, 0.31] to 0.68 [0.64, 0.71]. This range of correlations suggests that while several scales may be very well represented by the new CPAI items, some scales may not have been reproduced as closely. A closer examination of the pattern of correlations among the 11 new CPAI scales reveals that the new scales replicate the pattern of intercorrelations seen among the CPAI-2 scales as is illustrated by the shaded regions in Figure 2. Similar to the original set of emic scales, the new CPAI scales also show high

intercorrelations among related scales. Overall, the new CPAI scales show adequate convergence with the full-length CPAI-2 scales.

Correlations with the Big Five. Table 9 presents the correlations between BFI factors and the new CPAI and CPAI-2 scales. These correlations are also highlighted in Figures 4 and 5 to illustrate the differences in the patterns of correlations between 0the two sets of emic scales. These results show that most of the new CPAI scales adequately reflect the patterns of correlations between the CPAI-2 and BFI scales, and are similarly highly correlated with several BFI factors. This provides more evidence for the convergence of the new measure. Notably, the Discipline scale did not replicate these correlations very well with three BFI factors correlating inversely. This warrants further examination of these items.

Criterion-related validity. Table 10 presents the correlations between selected criterion scales and the new CPAI and CPAI-2 scales. These correlations are also highlighted in Figures 6 and 7 to illustrate the differences in the patterns of correlations. Again, we see that most of the scales mimic the correlations of their CPAI-2 counterparts' relationships with the OCB and CWB scales, suggesting good criterion-related validity and convergence for these work-related constructs. Correlations with the psychological wellbeing factors were also largely consistent, further suggesting decent convergence.

The Discipline scale again showed differential criterion-relatedness with some correlations seeming reversed when comparing the new CPAI and CPAI-2 scales. The Face and Traditionalism vs. Modernity scales also seemed to be functioning somewhat differently from their CPAI-2 equivalent scales.

These results overall suggest that the new CPAI items are adequately able to address and represent the construct space assessed by the CPAI-2 emic scales. The next section explores the scale

shortening methods, discussing the criteria used to select and omit items. The few issues mentioned above will be considered when assessing the items for retention in the final Short CPAI scale.

Shortening the Scales

Item Factor Analysis. Items were analyzed at the scale level using the *mirt* package to examine IRT discrimination and intercept parameters. IRT analyses mentioned here concern the full sample, however, we also checked for differential functioning among the items between the two groups. As such, analyses were additionally conducted at the group level and we noticed interesting differences between the two groups. These differences deemed not to produce drastically differential functioning of the new CPAI scales and are discussed further below.

Item statistics of all newly constructed CPAI items for the full sample are presented in Table 7. Figures 8 to 18 show the item trace plots for each scale. Note that these plots show two lines corresponding to the Chinese and US samples. Items were selected to be retained based on the information curves and information parameters such that items that were most highly correlated with the scale and seemed to represent high information were retained.

Using the IFA method, a total of 52 items were retained, shortening the inventory by 25 items. All scales have between 4-6 items as previously aimed, except for the Discipline scale which only had three viable items. Item characteristics and scale reliabilities for the shortened CPAI scale developed using the IFA approach are reported in Table 11. Scale and item information for the retained items are illustrated in figures 24-34.

Convergence with full-length scales. Correlation between scale scores computed using these 52 items and the equivalent CPAI-2 scales are reported in Table 12. These correlations are largely similar to those of the full set of new CPAI items and range from r [95% CI] = 0.09 [0.03, 0.15] to 0.64 [0.60, 0.67]. While the lowest correlation was among the Discipline scales at $r = 0.09$, all other scale correlations were $r > 0.3$ suggesting moderate convergence between the IFA-shortened scales

and the original CPAI-2 scales. Similar to the full set of new CPAI items, the shortened scales do not all highly correlate with their CPAI-2 counterparts, suggesting that some scales may not be replicating those concept spaces as effectively. These correlations are compared against the full scale correlations in Figure 19.

Correlations with the Big Five. Table 13 presents the correlations between BFI factors and the IFA-shortened new CPAI scales. These correlations are also highlighted in Figure 20 and 21, and are compared to the full scale correlations. The shortened scales exhibit similar patterns of correlations as the full scale, and resemble the CPAI-2 scale relationships with the BFI factors, providing additional support for the scale's similarity when shortened. However, the Discipline scale did not replicate as closely as the other scales.

Criterion-related validity. Table 14 presents the correlations between criterion scales and the shortened scales, which are also illustrate in Figures 22 and 23. Correlations are again largely similar to the full scale and the CPAI-2 equivalent scales, providing more evidence for convergence. While the Discipline scale did not reflect the full set of new items very closely, it did resemble the patterns of correlations between the original CPAI-2 scales and the criterion variables.

Genetic Algorithm. The algorithm was simulated to iterate across a series of item cost values. This approach produced several potential inventories of different lengths according to the item cost of each simulation, with the smallest item cost, 0.005 producing the longest shortened scale with 27 total items, while the largest item cost used, 0.09, produced an inventory with only 14 total items. The resulting cross-validation correlations of each simulation and matrix of the number of items retained by the GA for each scale are reported in Tables 15 and 16.

The solution produced with item cost fixed at 0.03 seemed to produce the most parsimonious set of items that minimized the number of items but maximized the cross-validation correlation.

While the lowest item cost of 0.005 also converged similarly, the shorter resulting scale was selected since the aim of this study was to develop a short scale. Since the shorter scale performs adequately, we decided to retain it.

The GA produced much fewer items than the IFA approach with the total item count for the retained inventory at only 20 items. Additionally reflecting the efficiency of the algorithmic approach, five scales resulted in single-item solutions. Table 17 reports the item characteristics and scale reliabilities for all retained items. Scale and item information plots for the retained items are illustrated in figures 35-40.

Although we initially proposed to select 4-6 items for each scale, in minimizing the redundancy of retained items, the GA selected far fewer items than expected. However, the resulting items are surprisingly reflective of the scales, for example, for the Social Sensitivity scale, the retained single item reads “I do not like working on team projects”. Similarly, the single item retained for the Renqing scale reads “I feel obligated to help people who helped me”. Both of these items are highly characteristic of the scale they are assessing, allowing us to provide human affirmation to the machine’s suggestion.

Convergence with full-length scales. Correlations between scale scores computed using these 20 items and the equivalent CPAI-2 scales are reported in Table 18. Again, the correlations are very similar to those of the full set of new CPAI items as well as the IFA-shortened set of items. Correlations range from r [95% CI] = 0.30 [0.24, 0.35] to 0.56 [0.51, 0.60].

This suggests moderate convergence between the GA-shortened scales and the original CPAI-2 scales. Observing that the patterns of correlations hold even when some scales were shortened to a single item (e.g., the Graciousness vs. Meanness scale which has a resulting r = 0.55) suggests that short analogues to the CPAI-2 scales could adequately replicate the construct space.

Correlations with the Big Five. Table 19 presents the correlations between BFI factors and the GA-shortened new CPAI scales. These correlations are also highlighted in Figure 20 and 21, and are compared to the full scale correlations as well as the original CPAI-2 correlations. The GA-shortened scales mostly exhibit similar patterns of correlations as the full scale, and resemble the direction and magnitude of most CPAI-2 scale relationships with the BFI factors. However, the Discipline scale seems to show inverse relationship with the factors in comparison to the full scale and the IFA approach.

These results provide additional support for the scale's similarity and convergence when shortened, and also highlight the low functioning of the Discipline scale, as observed previously.

Criterion-related validity. Table 20 presents the correlations between criterion scales and the shortened scales, which are also illustrated in Figures 22 and 23. Correlations are again largely similar to the full scale and the CPAI-2 equivalent scales, however, the Discipline and Face scales did not seem to replicate these patterns as closely. On the other hand, the existing similarity between the correlations given such a drastically shortened scale still provides promising evidence for the convergence of this short scale with the full set of items developed in this study as well as with the original CPAI-2 items.

Chapter 4: Discussion

In this study, we investigated two separate processes for shortening psychometric scales and generated short public-domain versions of CPAI-2 scales. Specifically, we explored a traditional item factor analytic method and a novel genetic algorithm to shorten analogues to emic traits assessed by the CPAI-2. Our focus on the CPAI-2 was motivated by its properties, i.e., its length and proprietary nature, in an attempt to develop short measures for the public domain that can be used to further test the cultural-specificity hypothesis of personality traits. Below we compare the relative advantages of each approach, discuss the implications of automated technology for scale development and the practical implications of these new short scales, and examine the limitations faced in this study.

Item Screening Processes

The genetic algorithm was by and large a much more efficient method of shortening the scale. Limiting the manual labor to the construction of the algorithm and specification of its parameters meant that none of the actual item selection was done by humans. Instead, the simulation tested iterative combinations of all items to select only those that maximized convergence with the original CPAI-2 scales and minimized redundancy of remaining items. As such, several scales were constructed with very few items that were able to assess the respective trait broadly.

On the other hand, the factor analytic approach allowed us to examine each item in its entirety, focus on its relationships with other items in the scale, and use several pieces of information to determine which items to retain. As human readers, we were also able to judge the qualitative item content for trait relevance. Although several quantitative aspects of the items were considered, the process was necessarily subjective and relied on knowledge of scale construction, past experience,

and expertise. The final short scale developed using this method was longer than the scale constructed by the GA, but shared several items identified by the GA, as reported in Table 21.

The clear advantage that the GA provides is the efficient use of time. This automated approach, introduced by Yarkoni (2010), is able to abbreviate a measure with relatively little loss of convergence with the original measure, and provides considerable time-saving benefits to researchers. This study provides further support for the flexibility and robustness of the GA. Furthermore, our simulation of the item costs illustrates the ability to balance the needs of brevity with fidelity as required by the researcher's aims.

At the same time, there were several solutions identified by the GA where some scales were not identified by any of the items constructed to measure the scale due to cross-loading and due to the GA's tendency to minimize redundant items. For example, at higher item costs than the retained solution, items designed to assess the Renqing, Social Sensitivity, and Discipline scales were not retained in the final solution as items from other scales were deemed sufficient enough to capture the variance in these scales. This reflects the highly pragmatic nature of the GA which is driven primarily by the need to produce a short, reflective measure rather than one that is constructed using conventional psychometric and content relevance criteria (Eisenbarth et al., 2015). This may be detrimental to some who want to develop short scales but do not want to lose the content relevance of items.

It is also difficult to assess the reliability of short scales designed to minimize redundancy using traditional methods like internal consistency because high internal consistency would necessarily imply existing redundancies among items (see: Yarkoni, 2010).

As such, the process of identifying the appropriate approach is highly dependent on the context and requirements.

While the item screening process certainly differed in terms of time, labor, and effort, there was little difference in the resulting scales' performance. We observed that both short scales function quite similarly to the full new CPAI measure, and were also very similar to the original CPAI-2 constructs in terms of their relationships within the greater nomological network. Both approaches served to produce functional and convergent short CPAI scales which can now be used in the public domain.

In most cases, a researcher intending to shorten a scale would have several reasons to do so, but would fundamentally rely on creating a short scale that is able to effectively mirror the longer version of the measure. If speed is the goal, there is no argument that the GA is a quick approach. After spending some time familiarizing oneself with the parameters and specifications of the model, it is as simple as tweaking the numbers and plugging in the data to receive a packaged short scale. It is almost as if it is a gift. The GA also optimizes fit with the original scale scores, which is often not an explicit consideration in traditional manual approaches.

However, this approach may not work in all scenarios. We saw in our study that some scales were not identified by any items written to measure that scale. This would be problematic in cases like the current one where one needs to have items that were specifically designed to assess a particular construct. As such, it is fundamentally necessary to be aware of the item's properties as a responsible researcher. One should not simply put all their faith into the black box of the algorithm and expect a complete and perfect set of results. Researchers should attempt to use time-saving automated approaches but must combine their efforts with their knowledge of psychometrics in order to ensure that items that are retained are still psychometrically sound.

Importantly, the GA is just one approach towards automated abbreviation of personality scales. There may be several alternative methods, such as ant colony optimization (cf. Schroeders et

al., 2016) that may produce equivalent or better results, and may differ in terms of suitability under different circumstances. As knowledgeable humans, we have the ability to judge and compare different models, and we must continue doing so in order to identify and improve existing automated approaches to scale construction.

Lastly, since the GA optimizes prediction of the long scale when constructing short measures, any version it creates would also replicate the shortcomings of the original scale (Schroeders et al., 2016). This is an important concern for researchers intending to use this automated method, and introduces the potential for exacerbating existing deficiencies rather than eliminating them. In our study, we discussed several psychometric limitations of the CPAI-2 and yet, we used its scale scores to construct the short versions. Schroeders and colleagues advocate against the use of a GA without optimization features for models with poor fit, and we tend to agree with this advise. As a responsible steward of scale construction, one must be aware of the nature and fit of their scales' items before tossing the data into a simulation. We recommend an approach that is more synergistic of human and machine efforts, avoiding total reliance on either end of the spectrum.

Practical Implications of the New CPAI Scales

We were first inspired to conduct this study in order to explore the potential for culturally-specific traits, as recognized by the developers of the CPAI-2 (F. M. Cheung, 2006). The CPAI-2 is difficult to investigate because of its length, proprietary status, and questionable psychometric properties as reported in past studies. We hoped to provide opportunities for further probing the properties of the CPAI-2 constructs by identifying public-domain analogues to its several scales. We identified existing measures that could serve as analogues for its etic scales and then developed our own set of items to assess its emic traits.

The newly created scales display properties similar to the original scale and ranged in level of convergence. Additionally, the convergence displayed by the short scales derived from the IFA

method and the GA method was similar to each other as well as to the convergence displayed by the full set of newly generated analogue items. Some scales had high convergence with the equivalent original CPAI-2 scales, e.g., for Family Orientation: r_{IFA} [95% CI] = 0.64 [0.60, 0.67], r_{GA} [95% CI] = 0.53 [0.49, 0.58]; while other scales displayed significant correlation but low convergence, e.g., for Discipline: r_{IFA} [95% CI] = 0.09 [0.03, 0.15], r_{GA} [95% CI] = 0.32 [0.26, 0.38].

Furthermore, the new scales were able to replicate the pattern of intercorrelations reflected by the original set of CPAI-2 scales (see Figure 2) as well as the patterns of correlations with the Big Five and the criterion scales (see Figures 19-22). We observed several large correlations among the newly created short-form scales. This pattern of intercorrelations is largely similar to the intercorrelated scales of the original CPAI-2. We observed that a majority of the newly produced scale, both as a full set as well as when shortened using the IFA and GA approaches, reflected the same direction and similar magnitudes of correlation as the original CPAI-2 scales. This provides good evidence for the similarity of most of the short analogue scales generated in this study. However, similar to the previous issue, the Discipline scale did not replicate the correlation patterns as well as the other scales. The Discipline scale in particular seems to have underperformed in contrast to the other scales generated in this study. Future research should explore the content of the newly generated items and ensure that the description of the Discipline construct described by the CPAI-2 is capture by these sets of items.

Although the convergent correlations may seem low, these are typically observed when creating short form scales. For example, Gosling and colleagues (2003) published the very brief five- and ten-item short forms of the 44-item BFI where they reported convergent correlations ranging from 0.48 to 0.87. Similarly, Yarkoni (2010) abbreviated several personality scales and reported mean convergent correlations ranging from 0.53 to 0.67. In another study, Donnellan et al. (2006)

shortened the 50-item IPIP-FFM inventory by selecting the 20 most discriminating items with convergent correlations ranging from 0.83 to as high as 0.94. Higher convergence is observed here when the smaller scale uses a subset of the larger scale, however, when comparing the 20-item Mini-IPIP with a five-factor analogue, the IPIP-NEO, we see convergent correlations ranging from 0.52 to 0.73 at the factor level, and from 0.05 to 0.74 at the scale level. However, in all these cases, we see that the pattern of relationships that the short versions have with other personality measures and with criterion measures is very similar to their larger, original counterparts[5]. This reflects the psychometric cost of using short measures which, when compared to the long versions, tend to be less reliable and more susceptible to variance in the sample, especially when assessing multi-faceted scales as opposed to broad factors (Gosling et al., 2003).

Given these observations, the results observed in this study are not surprising when also considering the extent of the abbreviation, from 341 items in the original measure to 52- and 20-item analogues, and the fact that the short form items were newly constructed public-domain analogues, instead of exact replicas. This then brings us to a discussion about the replicability established by these new scales. We do not intend for these scales to be a replacement for the use of the original CPAI-2 scales. On the contrary, we intend for these scales to serve as public-domain analogues to the sets of constructs assessed by the CPAI-2 scales. The items used in the short-form measures are not derived from the original CPAI-2 inventory but were generated solely for the purpose of representing the construct spaces identified by the CPAI-2 scales. We intend to provide assessments of the constructs, specifically the culturally-specific emic constructs identified by the CPAI-2, to motivate

[5] The authors would like to note two publications, Zhou et al. (2021) and Dong et al. (2020), that emerged after the beginning of the current project. These publications report the development of short-form versions of the CPAI-2 and an adolescent version of the CPAI. Notably, both publications largely report psychometric properties only at the factor level and not at the scale level, while neither publication reports convergence with the full original scale. Since we are considering convergence at the scale level as the main criteria for comparison and evaluation of short-form scales, we were not able to consider the scales developed in these publications within this report.

more research into the cultural specificity of personality. In its current form, the CPAI-2 remains the only widely-used inventory and model of Chinese personality traits, however, research into these traits is limited due to its proprietary nature. We hope that these scales will allow for more accessible research probing the potential for capturing true culturally-specific traits and nomological relationships of these constructs.

We also observed some interesting differences between the Chinese and American samples, although we did not specifically set out to explore statistical differences between the samples. First, we noticed that some scales tended to be more informative in the Chinese sample than in the US sample. In particular, two of the scales designed to assess these alleged emic traits, Graciousness vs Meanness and Traditionalism vs Modernity, tended to provide more information overall in the Chinese sample than in the US sample. This suggests potential differences between the samples in the understanding and personal relevance of constructs assessed. It is possible that the items assessing these two traits held together more for Chinese respondents than US respondents, suggesting a potentially greater sense of cohesiveness of those constructs for Chinese participants.

At the same time, we expected the Renqing scale to be more differentiated between the two samples, and yet this was not the case. It is possible that this trait's notion of reciprocal interactions, e.g., "If a friend gives me a gift, I have to give them a gift back at some point", is a more universally-recognizable concept than other constructs (such as Traditionalism vs. Modernity which would involve distinguishing between a predilection for cultural tradition and innovation).

Personality is considered to be generally stable and universally expressed (Briley & Tucker-Drob, 2014; Ferguson, 2010; McCrae et al., 2005). In fact, it is suggested that the stability of personality is difficult to be studied meaningfully without also examining the stability and change present in individuals' environments (Ardelt, 2000). This stands to reason that while personality traits

are stable, they are fundamentally associated with one's life and surroundings. Given this view, it is not at all surprising to consider that vast cultural differences may result in vast culturally specific ways of thinking and being. It is also not too far of an extension to propose that these culture-linked behaviors would be predicted by culture-linked personality traits as suggested by Cheung and colleagues (2006; 2013). Whether these cultural influences are linked to unique traits that represent culturally-specific sets of values, or are more representative of characteristic adaptations one develops alongside personality traits to thrive in a given environment, is still a topic of debate.

In order to continue investigating these constructs, the structure of the CPAI-2, and the potential for the cultural-specificity of personality, we call on researchers to use and continue refining the short-form scales developed in this study. We show that many of the short scales are able to replicate the construct space defined by the original scales and do relate meaningfully to relevant life criteria. As such, these scales may be following similar predictive paths as the construct of interest identified in the CPAI-2. To that end, some scales in our new, short measures were better able to assess those constructs than other scales. We encourage further iteration of these short analogue scales in order to better represent the construct definitions with a small set of publicly-available items.

Limitations

The results in this study validated the use of an automated platform for abbreviating personality measures. The GA was able to effectively shorten a 77-item scale to a 20-item scale, which is less than 10 times shorter than the original CPAI-2 which has a whopping 341 items. While we claim success in our mission, there are a few key limitations we would like to discuss.

First, regarding the shallow correlations between our new measure and the original CPAI-2, there are several differences between the two scales in terms of their construction and use. The CPAI-2 was constructed to specifically assess concepts familiar in Chinese cultures by developers

who are intimately familiar with Chinese cultural norms. Our scale on the other hand, was developed to mimic the constructs assessed by the emic scales in the CPAI-2 using the scale descriptors. In addition, our items were written primarily by graduate students of Chinese origin as well other students from non-Chinese backgrounds. As such, it is not entirely surprising that some scales show lower fit than others. Future research should strive to continue testing the concepts proposed by the CPAI-2 and may benefit from taking a more emic approach to its identification and item generation. In addition, our scales were prepared in English and then translated to Simplified Chinese by a team of graduate student researchers who, although are fluent in Mandarin Chinese, are not professionally trained translators. Future research should strive for indigenous development of measures designed to assess indigenous concepts while generating potential short scales.

Second, both scales use different rating formats. The original CPAI-2 uses a Yes/No response format while our newly constructed measure uses a traditional Likert-type agreement scale. It is possible that the differences in response formats and number of response options could influence the convergent validity correlation (Simms et al., 2019). Future research should continue probing these differences to assess the impact of response formats on different psychometric measures.

Third, as mentioned earlier, the GA suffers from the same pitfalls as the data provided to summarize scales. Given that the CPAI-2 suffers from several psychometric maladies, it is not surprising that we observed some solutions of the simulation where some scales were more identified by items designed to assess other scales. The highly intercorrelated scales pose some major threats to the validity of the algorithm and highlight two potential issues with the GA approach. The first is that the algorithm is fundamentally weakened by providing it with a highly intercorrelated measure. This makes it difficult for the GA to prioritize non-redundant items for each scale as we observed several items from other scales were identified as potential analogues in the short solutions. This leads us to

the second issue which is that the GA ignores several traditional psychometric standards for assessing item fit, namely that items are allowed to overlap across measures. While Yarkoni (2010) contends that the costs of allowing item overlap do not outweigh the benefits, future research should empirically probe this assertion and test for differences in shortened scale performance with and without scale overlap and across measures with a range of reliabilities.

Lastly, the CPAI-2 itself is a black box. We were given permission to use the scale for our research and were then provided with scale scores after data collection, however, the structure of the inventory and identification of items as belonging to each of the factors and scales is still a mystery. This makes it difficult to even begin exploring the issues discussed in this paper and we encourage future researchers to continue engaging with the public domain to make research findings more accessible and equitable.

Chapter 5: Conclusion

There is a universal desire to understand individual differences. We often look around us and wonder, why are they like this? Researchers have attempted to answer this question by proposing several theories of personalities, values, motivations, and other ways of conceptualizing one's place in the universe. In this study, we attempted to probe the nature of one measure that exists within the paradigm of culturally-specific personality traits. We identified public domain analogues for assessing the universal (etic) traits, and developed short-form analogue measures for the culture-specific (emic) traits. We also investigated two approaches for developing a short-form measure and discussed the relative merits of allowing an automaton to expedite this mission while using human awareness to judge its performance. To that end, we have delivered items that can be and must be further refined in order to continue testing personality phenomena and the cultural-specificity hypothesis. It is our hope that this method and these scales help encourage more research on the role that culture plays in personality.

References

Ackerman, P. L., & Kanfer, R. (2009). Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions. *Journal of Experimental Psychology: Applied*, *15*(2), 163–181. https://doi.org/10/dtw25j

Ardelt, M. (2000). Still Stable after All These Years? Personality Stability Theory Revisited. *Social Psychology Quarterly*, *63*(4), 392–405. https://doi.org/10.2307/2695848

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

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

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*(1/2), 9–30. https://doi.org/10.1111/1468-2389.00160

Berkel, V., & Kathryn, H. (2009). *The Relationship Between Personality, Coping Styles and Stress, Anxiety and Depression*. https://doi.org/10.26021/7463

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health, 6*. https://doi.org/10/gfsqzs

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance, 10*(2), 99–109. https://doi.org/10.1207/s15327043hup1002_3

Briley, D. A., & Tucker-Drob, E. M. (2014). Genetic and environmental continuity in personality development: A meta-analysis. *Psychological Bulletin, 140*(5), 1303–1331. https://doi.org/10.1037/a0037091

Burtăverde, V., de Raad, B., & Zanfirescu, A.-Ş. (2018). An emic-etic approach to personality assessment in predicting social adaptation, risky social behaviors, status striving and social affirmation. *Journal of Research in Personality, 76*, 113–123. https://doi.org/10/gfgnpr

Campbell, J. P., & Wiernik, B. M. (2015). The modeling and assessment of work performance. *Annual Review of Organizational Psychology and Organizational Behavior, 2*, 47–74. https://doi.org/10.1146/annurev-orgpsych-032414-111427

Cappelleri, J. C., Jason Lundy, J., & Hays, R. D. (2014). Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clinical Therapeutics, 36*(5), 648–662. https://doi.org/10/652

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Chen, S. X., Cheung, F. M., Bond, M. H., & Leung, J.-P. (2006). Going beyond self-esteem to predict life satisfaction: The Chinese case. *Asian Journal of Social Psychology*, *9*(1), 24–35. https://doi.org/10.1111/j.1467-839x.2006.00182.x

Cheung, F. M. (2006). A combined emic–etic approach to cross-cultural personality test development: The case of the CPAI. In Q. Jing, M. R. Rosenzweig, G. d'Ydewalle, H. Zhang, H.-C. Chen, & K. Zhang (Eds.), *Progress in psychological science around the world: Vol. 2. Social and applied issues* (pp. 91–104). Routledge. https://doi.org/10.4324/9781315793184-6

Cheung, F. M., Cheung, S. F., & Leung, F. (2008). Clinical utility of the Cross-Cultural (Chinese) Personality Assessment Inventory (CPAI-2) in the assessment of substance use disorders among Chinese men. *Psychological Assessment*, *20*(2), 103–113. https://doi.org/10/dzn83z

Cheung, F. M., Cheung, S. F., Leung, K., Ward, C., & Leong, F. (2003). The English version of the Chinese Personality Assessment Inventory. *Journal of Cross-Cultural Psychology*, *34*(4), 433–452. https://doi.org/10/c94sjp

Cheung, F. M., Cheung, S. F., Wada, S., & Zhang, J. (2003). Indigenous measures of personality assessment in Asian countries: A review. *Psychological Assessment*, *15*(3), 280–289. https://doi.org/10/c3cz88

Cheung, F. M., Kwong, J. Y. Y., & Zhang, J. (2003). Clinical validation of the Chinese Personality Assessment Inventory. *Psychological Assessment*, *15*(1), 89–100. https://doi.org/10.1037/1040-3590.15.1.89

Cheung, F. M., Leung, K., Fan, R. M., Song, W.-Z., Zhang, J.-X., & Zhang, J.-P. (1996). Development of the Chinese Personality Assessment Inventory. *Journal of Cross-Cultural Psychology*, *27*(2), 181–199. https://doi.org/10/djb3m6

Cheung, F. M., Leung, K., Zhang, J.-X., Sun, H.-F., Gan, Y.-Q., Song, W.-Z., & Xie, D. (2001). Indigenous Chinese personality constructs: Is the Five-Factor Model complete? *Journal of Cross-Cultural Psychology*, *32*(4), 407–433. https://doi.org/10/bgf78b

Cheung, F. M., Shu Fai Cheung, Jianxin Zhang, Leung, K., Leong, F., & Kuang Huiyeh. (2008). Relevance of Openness as a personality dimension in Chinese culture: Aspects of its cultural relevance. *Journal of Cross-Cultural Psychology*, *39*(1), 81–108. https://doi.org/10.1177/0022022107311968

Cheung, F. M., van de Vijver, F. J. R., & Leong, F. T. L. (2011). Toward a new approach to the study of personality in culture. *American Psychologist*, *66*(7), 593–603. https://doi.org/10.1037/a0022389

Cheung, S. F., Chan, W., & Cheung, F. M. (2007, July). *Applying the item response theory to personality assessment in organizational setting: A case study of the CPAI- 2 in the hotel industry*. 7th Conference of Asian Association of Social Psychology, Malaysia.

Cheung, S. F., Cheung, F. M., & Fan, W. (2013). *From Chinese to Cross-Cultural Personality Inventory A Combined Emic-Etic Approach to the Study of Personality in Culture*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199930449.001.0001/acprof-9780199930449-chapter-3

Cheung, S. F., Cheung, F. M., Howard, R., & Lim, Y.-H. (2006). Personality across the ethnic divide in Singapore: Are “Chinese Traits” uniquely Chinese? *Personality and Individual Differences*, *41*(3), 467–477. https://doi.org/10/bhdw6w

Chiaburu, D. S., Oh, I.-S., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, *96*(6), 1140–1166. https://doi.org/10.1037/a0024004

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64. https://doi.org/10/f5xfj8

Condon, D. M., Roney, E., & Revelle, W. (2017). A SAPA project update: On the structure of phrased self-report personality items. *Journal of Open Psychology Data*, *5*(1). https://doi.org/10.5334/jopd.32

Connelly, B. S., Ones, D. S., & Hülsheger, U. R. (2018). Personality in industrial, work and organizational psychology: Theory, measurement and application. In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *The SAGE handbook of industrial, work and organizational psychology* (2nd ed., Vol. 1, pp. 320–365). Sage. https://doi.org/10.4135/9781473914940.n13

Cook, R. D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society: Series B (Methodological)*, *48*(2), 133–155. https://doi.org/10.1111/j.2517-6161.1986.tb01398.x

Costa, P. T., & McCrae, R. R. (1992). *NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO-FFI)* [Professional manual]. Psychological Assessment Resources.

Credé, M., Harms, P. D., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology*, *102*(4), 874–888. https://doi.org/10.1037/a0027403

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19. https://doi.org/10.1016/j.jesp.2015.07.006

De Fruyt, F., & Wille, B. (2013). Cross-cultural issues in personality assessment. In *Handbook of personality at work* (pp. 333–355). Routledge. https://doi.org/10.4324/9780203526910

De Raad, B., Barelds, D. P. H., Levert, E., Ostendorf, F., Mlačić, B., Blas, L. D., Hřebíčková, M., Szirmák, Z., Szarota, P., Perugini, M., Church, A. T., & Katigbak, M. S. (2010). Only three factors of personality description are fully replicable across languages: A comparison of 14 trait taxonomies. *Journal of Personality and Social Psychology*, *98*(1), 160–173. https://doi.org/10.1037/a0017184

DeYoung, C. G. (2010). Personality neuroscience and the biology of traits. *Social and Personality Psychology Compass*, *4*(12), 1165–1180. https://doi.org/10.1111/j.1751-9004.2010.00327.x

DeYoung, C. G. (2015). Cybernetic big five theory. *Journal of Research in Personality*, *56*, 33–58. https://doi.org/10.1016/j.jrp.2014.07.004

DeYoung, C. G., & Krueger, R. F. (2018). A cybernetic theory of psychopathology. *Psychological Inquiry*, *29*(3), 117–138. https://doi.org/10.1080/1047840x.2018.1513680

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880

Dong, Y., Fan, W., Cheung, F. M., & Li, M. (2020). Development of a Short Form of the CPAI-A (Form B) with Rasch Analyses. *Journal of Applied Measurement*, *21*(4), 515–532.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP Scales: Tiny-yet-effective measures of the Big Five Factors of Personality. *Psychological Assessment, 18*(2), 192–203. https://doi.org/10.1037/1040-3590.18.2.192

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91*(1), 40–57. https://doi.org/10.1037/0021-9010.91.1.40

Eisenbarth, H., Lilienfeld, S. O., & Yarkoni, T. (2015). Using a genetic algorithm to abbreviate the Psychopathic Personality Inventory–Revised (PPI-R). *Psychological Assessment, 27*(1), 194–202. https://doi.org/10.1037/pas0000032

Ferguson, C. J. (2010). A meta-analysis of normal and disordered personality across the life span. *Journal of Personality and Social Psychology*, *98*(4), 659–667. https://doi.org/10.1037/a0018770

Fox, S., Spector, P. E., Goh, A., Bruursema, K., & Kessler, S. R. (2012). The deviant citizen: Measuring potential positive relations between counterproductive work behaviour and organizational citizenship behaviour. *Journal of Occupational and Organizational Psychology*, *85*(1), 199–220. https://doi.org/10.1111/j.2044-8325.2011.02032.x

Francis, L. J., & Pearson, P. R. (1988). The development of a short form of the JEPQ (JEPQ-S): Its use in measuring personality and religion. *Personality and Individual Differences*, *9*(5), 911–915. https://doi.org/10/c4tfp8

Gerlitz, J.-Y., & Schupp, J. (2005). *Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP: Dokumentation der Instrumententwicklung BFI-S auf Basis des SOEP-Pretests*

2005 (DIW Berlin Research Note No. 2005–4). Deutsches Institut für Wirtschaftsforschung. https://www.diw.de/documents/publicationen/73/43490/rn4.pdf

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

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg University Press. http://ipip.ori.org/A%20broad-bandwidth%20inventory.pdf

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

Goldman, S. H., & Raju, N. S. (1986). Recovery of One- and Two-Parameter Logistic Item Parameters: An Empirical Study. *Educational and Psychological Measurement*, *46*(1), 11–21. https://doi.org/10/ch6sr2

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*, 504–528. https://doi.org/10/cmz

Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality: Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, *46*(3), 355–359. https://doi.org/10.1016/j.jrp.2012.03.008

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/c9j35b

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*(2), 227–239. https://doi.org/10.1348/014466505x29657

Hofmans, J., Kuppens, P., & Allik, J. (2008). Is short in length short in content? An examination of the domain representation of the Ten Item Personality Inventory scales in Dutch language. *Personality and Individual Differences*, *45*(8), 750–755. https://doi.org/10/fv2k7v

Hough, L. M., & Ones, D. S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology* (Vol. 1, pp. 233–277). Sage. https://doi.org/10/bc67

Ion, A., Iliescu, D., Ilie, A., & Ispas, D. (2016). The emic–etic approach to personality measurement in personnel selection. *Personality and Individual Differences*, *97*, 55–60. https://doi.org/10.1016/j.paid.2016.02.082

ITC. (2005). *ITC guidelines for translating and adapting tests* (Technical Report ITC-G-TA-20140617). International Test Commission. https://www.intestcom.org/files/guideline_test_adaptation.pdf

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). Guilford Press.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). Guilford Press.

Johnson, J. A. (2014). Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality, 51*, 78–89. https://doi.org/10/bc99

Klein, D. N., Kotov, R., & Bufferd, S. J. (2011). Personality and Depression: Explanatory Models and Review of the Evidence. *Annual Review of Clinical Psychology*, *7*(1), 269–295. https://doi.org/10.1146/annurev-clinpsy-032210-104540

Kwong, J. Y. Y., & Cheung, F. M. (2003). Prediction of performance facets using specific personality traits in the Chinese context. *Journal of Vocational Behavior*, *63*(1), 99–110. https://doi.org/10.1016/s0001-8791(02)00021-0

Laher, S. (2015). Exploring the utility of the CPAI-2 in a South African sample: Implications for the FFM. *Personality and Individual Differences*, *81*, 67–75. https://doi.org/10.1016/j.paid.2014.12.010

Liao, W. (2005). *How personality and emotional intelligence affect work performance* [Master dissertation in Human Resources]. National Central University.

Lin, E. J.-L., & Church, A. T. (2004). Are indigenous Chinese personality dimensions culture-specific?: An investigation of the Chinese Personality Assessment Inventory in Chinese American and European American samples. *Journal of Cross-Cultural Psychology, 35*(5), 586–605. https://doi.org/10/c87ktv

Lin, H. P. (2004). *The impact of personality on creativity among R&D personnel: An example of the high- tech industry* [Master dissertation in Human Resources]. National Central University.

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49*(2), 433–442. https://doi.org/10/ggj9sz

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. https://doi.org/10/bhr39c

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*(1), 139–157. https://doi.org/10.1037/0022-3514.88.1.139

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

McCrae, R. R., Costa, P. T., del Pilar, G. H., & Parker, W. D. (1998). Cross-cultural assessment of the five-factor model: The revised NEO Personality Inventory. *Journal of Cross-Cultural Psychology, 29*(1), 171–188. https://doi.org/10.1177/0022022198291009

McCrae, R. R., Terracciano, A., & 78 members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88*(3), 547–561. https://doi.org/10.1037/0022-3514.88.3.547

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. https://doi.org/10.1037/a0028085

Ng, A., Fan, W., Cheung, F. M., Leong, F. T. L., & Cheung, S. F. (2012). The CPAI-2 as a culturally relevant personality measure in differentiating among academic major groups. *Journal of Career Assessment, 20*(2), 196–207. https://doi.org/10.1177/1069072711420857

Nikan, F., Asghari Jafarabadi, M., Mohammad-Alizadeh-Charandabi, S., & Mirghafourvand, M. (2018). Designation and psychometric properties of the Short Form Postpartum Quality of Life Questionnaire (SF-PQOL): An application of multidimensional item response theory and genetic algorithm. *Health Promotion Perspectives*, *8*(3), 215–224. https://doi.org/10/gd2gjq

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw Hill.

Oh, I.-S. (2009). *The Five Factor Model of personality and job performance in East Asia: A cross-cultural validity generalization study* [Doctoral dissertation, University of Iowa]. http://search.proquest.com/dissertations/docview/304903943/

Olaru, G., Witthöft, M., & Wilhelm, O. (2015). Methods matter: Testing competing models for designing short-scale Big-Five assessments. *Journal of Research in Personality*, *59*, 56–68. https://doi.org/10/f7x9xm

Ones, D. S., & Dilchert, S. (2013). Counterproductive work behaviors: Concepts, measurement, and nomological network. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology* (Vol. 1, pp. 643–659). American Psychological Association. https://doi.org/10.1037/14047-035

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, *81*(6), 660–679. https://doi.org/10.1037/0021-9010.81.6.660

Ones, D. S., Wiernik, B. M., Wilmot, M. P., & Kostal, J. W. (2016). Conceptual and methodological complexity of narrow trait measures in personality-outcome research: Better knowledge by

partitioning variance from multiple latent traits and measurement artifacts. *European Journal of Personality*, *30*(4), 319–321. https://doi.org/10.1002/per.2060

Orlando, M. (2004, June). *Critical Issues to Address when Applying Item Response Theory (IRT) Models*. Drug Information Association meeting, Bethesda, MD.

Orlando, M., & Marshall, G. N. (2002). Differential item functioning in a Spanish translation of the PTSD checklist: Detection and evaluation of impact. *Psychological Assessment*, *14*(1), 50–59.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, *5*(3), 343–355. https://doi.org/10.1037//1082-989x.5.3.343

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*(1), 401–421. https://doi.org/10/d7hjz3

Porter, C., Woo, S. E., & Tak, J. (2016). Developing and validating short form protean and boundaryless career attitudes scales. *Journal of Career Assessment*, *24*(1), 162–181. https://doi.org/10/gd8c79

Pritikin, J. N., Hunter, M. D., & Boker, S. M. (2014). Modular open-source software for Item Factor Analysis. *Educational and Psychological Measurement*. Advance online publication. https://doi.org/10.1177/0013164414554615

R Core Team. (2020). *R: A language and environment for statistical computing* (4.0.2) [Computer software]. R Foundation for Statistical Computing. http://www.r-project.org/

Raborn, A. W., Leite, W. L., & Marcoulides, K. M. (2020). A Comparison of Metaheuristic Optimization Algorithms for Scale Short-Form Development. *Educational and Psychological Measurement*, *80*(5), 910–931. https://doi.org/10/ghhmwq

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Revelle, W. (2019). *psych: Procedures for psychological, psychometric, and personality research* (1.9.12) [R package]. Northwestern University. https://CRAN.R-project.org/package=psych

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345. https://doi.org/10/cq42fv

Rolland, J.-P. (2002). The cross-cultural generalizability of the five-factor model of personality. In R. R. McCrae & J. Allik (Eds.), *The five-factor model of personality across cultures* (pp. 7–28). Springer US. https://doi.org/10.1007/978-1-4615-0763-5_2

Şahin, A., & Anıl, D. (2016). The effects of test length and sample size on item parameters in item response theory. *Educational Sciences: Theory & Practice*, *17*(1). https://doi.org/10/gf2zt4

Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, *82*(1), 30–43. https://doi.org/10.1037/0021-9010.82.1.30

Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, *76*(3), 323–346. https://doi.org/10.1348/096317903769647201

Samejima, F. (1997). Graded Response Model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 85–100). Springer New York. https://doi.org/10.1007/978-1-4757-2691-6_5

Sandy, C. J., Gosling, S. D., & Koelkebeck, T. (2014). Psychometric comparison of automated versus rational methods of scale abbreviation. *Journal of Individual Differences, 35*(4), 221–235. https://doi.org/10.1027/1614-0001/a000144

Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar big-five markers. *Journal of Personality Assessment, 63*(3), 506–516. https://doi.org/10/cfq

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

Schmitt, D. P. (2004). The Big Five related to risky sexual behaviour across 10 world regions: Differential personality associations of sexual promiscuity and relationship infidelity. *European Journal of Personality, 18*(4), 301–319. https://doi.org/10.1002/per.520

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schroeders, U., Wilhelm, O., & Olaru, G. (2016). Meta-heuristics in short scale construction: Ant colony optimization and genetic algorithm. *PLOS ONE, 11*(11), e0167110. https://doi.org/10.1371/journal.pone.0167110

Simms, L., Zelazny, K., Williams, T., & Bernstein, L. (2019). Does the Number of Response Options Matter? Psychometric Perspectives Using Personality Questionnaire Data. *Psychological Assessment, 31*. https://doi.org/10.1037/pas0000648

Soto, C. J. (2019). How 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

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior*, *68*(3), 446–460. https://doi.org/10.1016/j.jvb.2005.10.005

Stanek, K. C., & Ones, D. S. (2018). Taxonomies and compendia of cognitive ability and personality measures relevant to industrial, work, and organizational psychology. In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *The SAGE handbook of industrial, work and organizational psychology: Vol. 1: Personnel psychology and employee performance* (2nd ed., pp. 366–407). Sage. https://doi.org/10.4135/9781473914940.n14

Stankov, L. (2011). Individual, country and societal cluster differences on measures of personality, attitudes, values, and social norms. *Learning and Individual Differences*, *21*(1), 55–66. https://doi.org/10/ffzt9p

Tabachnick, B. G., & Fidell, L. S. (2018). *Using Multivariate Statistics* (7 edition). Pearson.

Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, *51*(4), 567–577. https://doi.org/10/d2m8zd

Tyler, G. P., & Newcombe, P. A. (2006). Relationship between work performance and personality traits in Hong Kong organizational settings. *International Journal of Selection and Assessment*, *14*(1), 37–50. https://doi.org/10.1111/j.1468-2389.2006.00332.x

van Aarde, N., Meiring, D., & Wiernik, B. M. (2017). The validity of the Big Five personality traits for job performance: Meta-analyses of South African studies. *International Journal of Selection and Assessment*, *25*(3), 223–239. https://doi.org/10.1111/ijsa.12175

Wiernik, B. M., Bornovalova, M. A., Stark, S. E., & Ones, D. S. (2019). Constructs versus measures in personality and other domains: What distinguishes normal and clinical? *Industrial and Organizational Psychology*, *12*(2), 157–162. https://doi.org/10.1017/iop.2019.31

Wiernik, B. M., & Dahlke, J. A. (2020). Obtaining unbiased results in meta-analysis: The importance of correcting for statistical artifacts. *Advances in Methods and Practices in Psychological Science*, *3*(1), 94–123. https://doi.org/10.1177/2515245919885611

Willighagen, E. (2005). *genalg: R Based Genetic Algorithm* (R package version 0.1. 2005.) [Computer software].

Woo, S. E., Chernyshenko, O. S., Longley, A., Zhang, Z.-X., Chiu, C.-Y., & Stark, S. E. (2014). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of Personality Assessment*, *96*(1), 29–45. https://doi.org/10.1080/00223891.2013.806328

Xie, Q., Fan, W., Wong, P., & Cheung, F. M. (2016). Personality and parenting style as predictors of life satisfaction among Chinese secondary students. *The Asia-Pacific Education Researcher*, *25*(3), 423–432. https://doi.org/10/gfwm2n

Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., Ostendorf, F., Angleitner, A., Riemann, R., Spinath, F. M., Livesley, W. J., & Jang, K. L. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*(6), 987–998. https://doi.org/10.1037/0022-3514.90.6.987

Yarkoni, T. (2010). The abbreviation of personality, or how to measure 200 personality scales with 200 items. *Journal of Research in Personality*, *44*(2), 180–198. https://doi.org/10/bprznw

Zhou, M., Huang, D., Ren, F., Fan, W., Mu, W., Li, F., Zhang, J., & Cheung, F. M. (2021). Short Forms of the Cross-Cultural (Chinese) Personality Assessment Inventory: Reliability, Validity, and Measurement Invariance Across Gender. *Frontiers in Psychology*, *12*, 709032. https://doi.org/10.3389/fpsyg.2021.709032

Appendix A:
Tables

Table 1. Four factors and 28 scales from the CPAI-2.

Factor	Scale	Emic (M) or Etic (T)	No. of items	Description
Social Potency	Novelty (NOV)	T	10	Extent to which one likes trying new things and facing new challenges. **High:** Open to trying new things and tackling new challenges. **Low:** Resists new ideas and experiences, and dislikes abstract concepts. *Sample item: I do not like stable jobs; instead I like challenges.*
	Diversity (DIV)	T	10	Degree to which one likes trying out new ways of handling tasks and exposing self to diverse experiences. **High:** Enjoys exposure to diverse experiences. **Low:** Dislikes new things. *Sample item: I am not very interested in things unrelated to my job. (Recode)*
	Divergent Thinking (DIT)	T	10	Extent to which one can deal with issues or problems from various perspectives. **High:** Greater tendency to deal with a problem by considering a variety of approaches. **Low:** Tends to stick to familiar concepts and approaches to solving problems. *Sample item: I always examine a particular issue from many different angles.*

Table 1. (Continued)

Factor	Scale	Emic (M) or Etic (T)	No. of items	Description
Social Potency	Leadership (LEA)	T	10	One's ability to take the initiative to lead, influence others, and make decisions in a group. **High:** Greater ability and motivation to influence others and lead decision making. **Low:** Lesser motivation to lead and greater tendency to follow. *Sample item: When several people are working together on something and there is no leader around, I will take over.*
	Logical versus Affective Orientation (L-A)	T	10	Extent of being objective or subjective in thinking and behavior. **High:** Objective, logical and analytic in thinking and behavior **Low:** Subjective sentimental and intuitive orientation. *Sample item: Before I make a decision, I will always analyze all the pros and cons.*
	Aesthetics (AES)	T	10	Degree to which one values and enjoys the beauty, arts and music in life. **High:** Greater tendency to enjoy the arts. **Low:** Lesser tendency to enjoy the arts. *Sample item: I think most poems are uninteresting and difficult to understand. (Recode)*
	Extraversion versus Introversion (E-I) (−)	T	10	One's social orientation and styles of interaction. **High:** Sociable and socially comfortable **Low:** Prefer to be quiet and solitary. *Sample item: I am very talkative when I am with a group of people.*
	Enterprise (ENT)	T	10	Extent to which one is prepared to explore the unbeaten paths and dare to take risks. **High:** Greater tendency to take risks and try novel things. **Low:** Greater tendency to stick to tried-and-tested methods. *Sample item: I do not like to take part in activities where I have to compete in public. (Recode)*

Table 1. (Continued)

Factor	Scale	Emic (M) or Etic (T)	No. of items	Description
Dependability	Responsibility (RES)	T	10	Extent to which one is dedicated, persistent, and can be relied upon to carry out tasks and achieve aims. **High:** Reliable and dedicated to goals. **Low:** Unreliable and often fails to complete necessary tasks. *Sample item: Often I start doing something else without having finished what I was doing before. (Recode)*
	Emotionality (EMO) (−)	T	10	Degree to which one is emotionally stable, in control of emotions. **High:** Deals well with stress and can control emotions well. **Low:** Experiences a lot of stress and drastic shifts in emotions. *Sample item: Sometimes I feel miserable for no reason.*
	Inferiority versus Self-acceptance (I-S) (−)	T	18	One's self-assurance and confidence. **High:** Poor self-esteem. **Low:** Self-confident. *Sample item: Even if I have already made a choice, I would easily regret and reverse it.*
	Practical Mindedness (PRA)	T	12	Extent to which one is realistic and pragmatic and one's focus on substance rather than form. **High:** More realistic and greater focus on substance. **Low:** Less practical in approach with a greater focus on style and form. *Sample item: I often day-dream. (Recode)*
	Optimism versus Pessimism (O-P)	T	10	Degree of one's positive or negative outlook on life and matters, and one's likelihood of worrying excessively or being critical of others. **High:** Energetic and positive outlook. **Low:** Holds grievances, low spirited. *Sample item: Even when I am in trouble, I still take a positive approach in handling the problems.*

Table 1. (Continued)

Factor	Scale	Emic (M) or Etic (T)	No. of items	Description
Dependability	Meticulousness (MET)	T	10	Degree of one's care for the quality of work products and attention to details. **High:** Cautious, orderly, and pays attention to details. **Low:** Does not expend much care for quality of work or fine details. *Sample item: I can remember very clearly what I have lent to and borrowed from others.*
	Face (FAC) (−)	M	11	One's concern for maintaining a proper reputation and image in social interactions. Focuses excessive attention to social recognition, concern for maintaining self-respect in social relationships. Face is a dominant concept in interpreting and regulating social behavior in Chinese culture. **High:** Places a lot of concern for one's reputation and image. **Low:** Does not care much about reputation or social-image. *Sample item: Even though I know I am wrong at times, I am not willing to admit it in public.*
	Internal versus External Locus of Control (I-E) (−)	T	10	Extent of attributing the causes of one's experience or events that happen to them. **High:** Attributes to internal factors in explaining success and failure **Low:** Attributes to external factors, luck and fate. *Sample item: Often I feel I have no control over what is happening to me. (Recode)*
	Family Orientation (FAM)	M	10	One's sense of family solidarity and responsibility. These family ties provide emotional and financial security and support, especially in Asian or collectivistic cultures. **High:** Value family bonding and have strong family ties. **Low:** Cares very little about responsibility to family and family solidarity. *Sample item: Usually I prefer to be with my intimate friends rather than my family. (Recode)*

Table 1. (Continued)

Factor	Scale	Emic (M) or Etic (T)	No. of items	Description
Accommodation	Defensiveness (Ah-Q Mentality) (DEF)	M	10	One's use of defense mechanisms, such as self-protective rationalization, externalization of blame, self-enhancement, and belittling of others' achievements. A mild degree is accepted as a healthy protective mechanism against defeatism and depression. **High:** Uses protective defense mechanisms to feel better about oneself. **Low:** Does not engage in a great deal defensive behaviors or thoughts. *Sample item: If other people do not invite me to their gathering, I see it as their own loss.*
	Graciousness versus Meanness (G-M)	M	10	Extent to which one is kind and broad-minded in their dealings with others. **High:** Bears no grudges, treat others leniently. **Low:** Overly critical of others, retaliatory, and calculating. *Sample item: I am sure if I were to encounter a misfortune, some people would take pleasure in it. (Recode)*
	Interpersonal Tolerance (IN T)	T	10	Extent to which one accepts diversity and tolerates differences in people. **High:** Accepts and tolerates people from all walks of life. **Low:** Intolerant of social diversity. *Sample item: I find it hard to accept a person whose personality is opposite to me. (Recode)*
	Self vs. Social Orientation (S-S)	M	10	Degree of enthusiasm for teamwork and willingness to contribute to the collective over the individual goals. **High:** Independent and unwilling to join cooperative activities **Low:** Collectivistic orientation, a team player. *Sample item: I feel happier when I am by myself.*
	Veraciousness versus Slickness (V-S)	T	10	One's reliability and consistency of their truthfulness. **High:** Truthful, adhere to principles **Low:** Boastful, suave, and superficial. *Sample item: I often flatter others in order to achieve my own goal. (Recode)*

Table 1. (Continued)

Factor	Scale	Emic (M) or Etic (T)	No. of items	Description
Interpersonal Relatedness	Traditionalism versus Modernity (T-M) (−)	M	15	The degree of individual modernization as an indication of one's responses to societal modernization. This scale covers attitudes toward traditional cultural beliefs and values in the areas of family relationships, filial piety, social rituals, and chastity that are emphasized in Chinese societies. **High:** Endorses traditional beliefs, customs, and values **Low:** Challenges traditional ideas, endorse individual freedom. *Sample item: If a dispute cannot be resolved, a family elder should be invited to act as an arbiter to uphold justice.*
	Renqing (Relationship Orientation) (REN)	M	12	One's adherence to the cultural norms of reciprocal interactions such as courtesy rituals, exchanging resources, maintaining and utilizing useful ties, and nepotism. **High:** Endorses cultural norms of courtesy and maintains useful ties. **Low:** Challenges cultural relationship norms and ideologies. *Sample item: If a friend or relative was hospitalized, I would definitely go visit him/her.*
	Social Sensitivity (SOC)	M	11	The extent to which individuals is empathic and sensitive to how others feel. **High:** Sensitive to how others feel and react. **Low:** Has little interest in other people's feelings. *Sample item: When I see others feeling distressed, I am easily affected too.*
	Discipline (DIS) (−)	M	11	How rigid and disciplined one is as opposed to being adaptable, flexible, and carefree. **High:** Rigid and disciplined. **Low:** Flexible and adaptable. *Sample item: Rules and laws should be strictly enforced and should be without exception.*

Table 1. (Continued)

Factor	Scale	Emic (M) or Etic (T)	No. of items	Description
Interpersonal Relatedness	Harmony (HAR)	M	14	The degree of one's inner peace of mind, contentment, and interpersonal relations with others. The avoidance of conflict and maintenance of equilibrium are considered important virtues in Asian traditions and collectivistic cultures. **High:** Maintains inner peace and contentment and avoids conflict and competition. **Low:** Often gets into conflict with others and faces a great deal of inner turmoil. *Sample item: I easily get into conflict with other people. (Recode)*
	Thrift versus Extravagance (T-E)	M	10	One's tendency to save rather than waste and one's carefulness in spending. Thrift is one of the basic traditional Confucian Chinese values, and the characteristic of thrift versus extravagance is an indicator of the social response to rapid economic development and increasing materialism. **High:** Endorses traditional value of frugality. **Low:** Endorses materialistic tendencies and high consumption. *Sample item: Even when I have new clothes, I continue to wear something old and save the new ones for important occasions.*

Note. Adapted from F. M. Cheung et al. (2008) and S. F. Cheung et al (2013).

Table 2. Patterns of correlations of CPAI-2 etic scales with NEO PI-R scales.

CPAI factor	CPAI scale	Patterns of correlations with NEO PI-R scales				
		C	A	N	O	E
Social potency	Novelty				+	
	Diversity				+	
	Divergent Thinking				+	
	Leadership				+	+
	Logical versus Affective Orientation				+	
	Aesthetics				+	
	Extraversion vs. Introversion					+
	Enterprise			−	+	
Dependability	Responsibility	+				
	Emotionality			+		
	Inferiority vs. Self-Acceptance			+		
	Practical Mindedness	+		−		
	Optimism vs. Pessimism			−		
	Meticulousness	+				
	Internal vs. External Locus of Control			−		
Accommodation	Interpersonal Tolerance		+			
	Voraciousness vs. Slickness		+			

Table 3. Public domain analogues identified for the CPAI-2 etic scales.

CPAI factor	CPAI etic scale	Description	Analogous constructs in Stanek and Ones' (2018) constructs	Analogue public domain or freely available scales
Social potency	Novelty	Extent to which one likes trying new things and facing new challenges. **High:** Open to trying new things and tackling new challenges. **Low:** Resists new ideas and experiences, and dislikes abstract concepts. ***Sample item:*** *I do not like stable jobs; instead I like challenges.*	Openness -- Variety Seeking	IPIP HEXACO O: Inquisitiveness
	Diversity	Degree to which one likes trying out new ways of handling tasks and exposing self to diverse experiences. **High:** Enjoys exposure to diverse experiences. **Low:** Dislikes new things. ***Sample item:*** *I am not very interested in things unrelated to my job. (Recode)*		IPIP NEO O4: Adventurousness
	Divergent thinking	Extent to which one can deal with issues or problems from various perspectives. **High:** Greater tendency to deal with a problem by considering a variety of approaches. **Low:** Tends to stick to familiar concepts and approaches to solving problems. ***Sample item:*** *I always examine a particular issue from many different angles.*	Openness -- Ideas	IPIP NEO O5: Intellect
	Leadership	One's ability to take the initiative to lead, influence others, and make decisions in a group. **High:** Greater ability and motivation to influence others and lead decision making. **Low:** Lesser motivation to lead and greater tendency to follow. ***Sample item:*** *When several people are working together on something and there is no leader around, I will take over.*	Extraversion -- Dominance	IPIP HPI: Leadership HIC
	Logical versus Affective Orientation	Extent of being objective or subjective in thinking and behavior. **High:** Objective, logical and analytic in thinking and behavior **Low:** Subjective sentimental and intuitive orientation. ***Sample item:*** *Before I make a decision, I will always analyze all the pros and cons.*		IPIP Emotion-based Decision-making (Barchard, 2001)

Table 3. (Continued)

CPAI factor	CPAI etic scale	Description	Analogous constructs in Stanek and Ones' (2018) constructs	Analogue public domain or freely available scales
Social potency	Aesthetics	Degree to which one values and enjoys the beauty, arts and music in life. **High:** Greater tendency to enjoy the arts. **Low:** Lesser tendency to enjoy the arts. ***Sample item:*** *I think most poems are uninteresting and difficult to understand. (Recode)*	Openness -- Aesthetics	IPIP NEO O2: Artistic Interests
	Extraversion vs. Introversion	One's social orientation and styles of interaction. **High:** Sociable and socially comfortable **Low:** Prefer to be quiet and solitary. ***Sample item:*** *I am very talkative when I am with a group of people.*	Extraversion	IPIP HEXACO X: Sociability
	Enterprise	Extent to which one is prepared to explore the unbeaten paths and dare to take risks. **High:** Greater tendency to take risks and try novel things. **Low:** Greater tendency to stick to tried-and-tested methods. ***Sample item:*** *I do not like to take part in activities where I have to compete in public. (Recode)*	compound - Risk Taking	IPIP Sensation-seeking facets (Hoyle et al., 2002): Calculated thrill-seeking
Dependability	Responsibility	Extent to which one is dedicated, persistent, and can be relied upon to carry out tasks and achieve aims. **High:** Reliable and dedicated to goals. **Low:** Unreliable and often fails to complete necessary tasks. ***Sample item:*** *Often I start doing something else without having finished what I was doing before. (Recode)*	Conscientiousness -- Dependability	IPIP NEO C5: Self-Discipline
	Emotionality	Degree to which one is emotionally stable, in control of emotions. **High:** Deals well with stress and can control emotions well. **Low:** Experiences a lot of stress and drastic shifts in emotions. ***Sample item:*** *Sometimes I feel miserable for no reason.*	compound - Self Control	IPIP Big 5 Factor IV (Emotional Stability)
	Inferiority vs. Self-acceptance	One's self-assurance and confidence. **High:** Poor self-esteem. **Low:** Self-confident. ***Sample item:*** *Even if I have already made a choice, I would easily regret and reverse it.*	compound - Self Esteem	IPIP NEO N4: Self-Consciousness

Table 3. (Continued)

CPAI factor	CPAI etic scale	Description	Analogous constructs in Stanek and Ones' (2018) constructs	Analogue public domain or freely available scales
Dependability	Practical Mindedness	Extent to which one is realistic and pragmatic and one's focus on substance rather than form. **High:** More realistic and greater focus on substance. **Low:** Less practical in approach with a greater focus on style and form. ***Sample item:*** *I often day-dream. (Recode)*	compound - Pragmatic	Items selected from IPIP AB5C: Cautiousness, NEO O1: Imagination, and 16PF M: Imagination scales.*
	Optimism vs. Pessimism	Degree of one's positive or negative outlook on life and matters, and one's likelihood of worrying excessively or being critical of others. **High:** Energetic and positive outlook. **Low:** Holds grievances, low spirited. ***Sample item:*** *Even when I am in trouble, I still take a positive approach in handling the problems.*	Neuroticism -- Negative Affect	IPIP CPI Optimism (Well-being [Wb])
	Meticulousness	Degree of one's care for the quality of work products and attention to details. **High:** Cautious, orderly, and pays attention to details. **Low:** Does not expend much care for quality of work or fine details. ***Sample item:*** *I can remember very clearly what I have lent to and borrowed from others.*	Conscientiousness - Orderliness	IPIP NEO C6: Cautiousness
	Internal vs. external locus of control	Extent of attributing the causes of one's experience or events that happen to them. **High:** Attributes to internal factors in explaining success and failure **Low:** Attributes to external factors, luck and fate. ***Sample item:*** *Often I feel I have no control over what is happening to me. (Recode)*	compound - Locus of Control	IPIP Locus of Control (Levenson, 1981): Total

Table 3. (Continued)

CPAI factor	CPAI etic scale	Description	Analogous constructs in Stanek and Ones' (2018) constructs	Analogue public domain or freely available scales
Accommodation	Interpersonal tolerance	Extent to which one accepts diversity and tolerates differences in people. **High:** Accepts and tolerates people from all walks of life. **Low:** Intolerant of social diversity. ***Sample item:*** *I find it hard to accept a person whose personality is opposite to me. (Recode)*	compound - Tolerance	IPIP JPI Tolerance [Tol]
	Veraciousness vs. slickness	One's reliability and consistency of their truthfulness. **High:** Truthful, adhere to principles **Low:** Boastful, suave, and superficial. ***Sample item:*** *I often flatter others in order to achieve my own goal. (Recode)*	NEO A2: Straightforwardness	IPIP NEO A2: Morality

Note. * IPIP items selected to assess the CPAI-2 Practical Mindedness scale - "Love to daydream.", IPIP 16PF: Imagination; "Do things by the book.", IPIP 16PF: Imagination; "Seldom get lost in thought.", IPIP 16PF: Imagination; "Have a vivid imagination.", IPIP AB5C: Imagination; "Rarely look for a deeper meaning in things.", IPIP MPQ: Imagination; "Indulge in my fantasies.", IPIP NEO: Imagination; "Sometimes have fantasies that are overwhelming.", IPIP CAT-PD: Fantasy Proneness; "Purchase only practical things.", IPIP AB5C: Cautiousness.

Table 4. Mean convergent correlations from the Genetic Algorithm simulation at different sample sizes.

Sample Size	Mean	SD	Range		95% CI for mean	
			Lower Bound	Upper Bound	Lower Bound	Upper Bound
100	0.655	0.066	0.500	0.737	0.642	0.668
250	0.650	0.052	0.453	0.736	0.637	0.663
375	0.645	0.046	0.434	0.732	0.633	0.658
500	0.642	0.043	0.431	0.731	0.629	0.655
750	0.638	0.042	0.416	0.729	0.626	0.651
1000	0.637	0.041	0.409	0.731	0.624	0.650

Table 5. Sample characteristics.

Characteristics		Overall sample		China		USA	
		N	%	N	%	n	%
Total sample size		918	100.00	313	34.10	605	65.90
Gender							
	Man	274	29.85	76	24.28	198	32.73
	Woman	481	52.40	85	27.16	396	65.45
	Nonbinary/ Genderqueer	7	0.76	0	0.00	7	1.16
	Transgender Male*	1	0.11	0	0.00	1	0.17
	Did not disclose	155	16.88	152	48.56	3	0.50
Race (US only)							
	American Indian or Alaska Native	1	0.17	-	-	1	0.17
	Asian	56	9.26	-	-	56	9.26
	Black or African American	62	10.25	-	-	62	10.25
	Hispanic or Latino	68	11.24	-	-	68	11.24
	Middle Eastern	4	0.66			4	0.66
	White	333	55.04	-	-	333	55.04
	Mixed Race	73	12.07	-	-	73	12.07
	Haitian*	1	0.17			1	0.17
	Caucasian*	1	0.17			1	0.17
	Did not disclose	6	0.99	-	-	6	0.99
Sexual Orientation							
	Completely Heterosexual	560	61.00	154	49.20	406	67.11
	Primarily Heterosexual	82	8.90	5	1.60	77	12.73
	Bisexual/Pansexual	87	9.50	1	0.32	86	14.21
	Primarily Homosexual	14	1.50	0	0.00	14	2.31
	Completely Homosexual	14	1.50	1	0.32	13	2.15
	Asexual	6	0.70	0	0.00	6	0.99
	Did not disclose	155	16.90	152	48.56	3	0.50
Education							
	Less than high school	2	0.20	1	0.32	1	0.17
	Some High School	1	0.10	0	0.00	1	0.17
	High School Degree	57	6.20	3	0.96	54	8.93
	Some College	328	35.70	6	1.92	322	53.22
	Bachelor's Degree	299	32.60	123	39.30	176	29.09
	Some Graduate School	8	0.90	2	0.64	6	0.99
	Master's Degree	64	7.00	23	7.35	41	6.78
	Doctoral Degree	4	0.40	3	0.96	1	0.17
	Did not disclose	155	16.90	152	48.56	3	0.50

Note. * indicates individuals who chose to self-identify as described.

Table 6. Scale means, standard deviations, and Cronbach's alpha for all measures in the overall sample, as well as the Chinese and US samples.

Scale	Factor	Overall Sample						
		N	M	SD	SE	α	95% CI L	95% CI U
New CPAI	FAC	918	20.96	3.85	0.13	0.46	0.40	0.51
	FAM	918	38.63	6.97	0.23	0.82	0.81	0.84
	DEF	918	19.74	3.93	0.13	0.48	0.43	0.53
	GM	918	19.31	3.21	0.11	0.32	0.25	0.38
	SS	918	17.74	5.50	0.18	0.78	0.76	0.80
	TM	918	18.92	4.79	0.16	0.75	0.73	0.78
	REN	918	29.12	3.66	0.12	0.69	0.66	0.72
	SOC	918	26.44	4.15	0.14	0.68	0.65	0.71
	DIS	918	14.17	3.20	0.11	0.49	0.44	0.55
	HAR	918	18.43	3.71	0.12	0.52	0.48	0.57
	TE	918	25.57	4.15	0.14	0.64	0.60	0.67
BFI	Conscientiousness	386	47.06	9.22	0.47	0.91	0.90	0.92
	Agreeableness	386	47.34	7.32	0.37	0.82	0.80	0.84
	Neuroticism	386	30.30	10.68	0.54	0.92	0.91	0.93
	Openness	386	47.84	8.08	0.41	0.88	0.87	0.90
	Extraversion	386	40.76	9.72	0.49	0.89	0.88	0.90
OCB	Total score	529	98.15	23.59	1.03	0.93	0.93	0.94
	Organizational	529	17.91	4.78	0.21	-	-	-
	People-directed	529	18.70	4.70	0.20	-	-	-
CWB	Total score	292	117.09	32.47	1.90	0.96	0.95	0.96
	Organizational	292	32.08	9.28	0.54	-	-	-
	People-directed	292	26.47	7.66	0.45	-	-	-
DASS-21	Depression	529	9.35	6.73	0.29	0.96	0.95	0.96
	Anxiety	529	9.24	5.46	0.24	0.91	0.90	0.91
	Stress	529	10.82	5.83	0.25	0.91	0.90	0.92
CPAI-2	FAC	918	5.67	2.31	0.08	-	-	-
	FAM	918	6.59	2.75	0.09	-	-	-
	DEF	918	2.50	2.15	0.07	-	-	-
	GM	918	6.69	2.47	0.08	-	-	-
	SS	918	5.38	2.15	0.07	-	-	-
	TM	918	4.95	3.07	0.10	-	-	-
	REN	918	8.56	2.00	0.07	-	-	-
	SOC	918	8.59	2.00	0.07	-	-	-
	DIS	918	6.14	2.29	0.08	-	-	-
	HAR	918	10.96	2.21	0.07	-	-	-
	TE	918	6.20	1.76	0.06	-	-	-

Table 6. (Continued)

Scale	Factor	China						
New CPAI	FAC	313	20.48	3.45	0.20	0.48	0.40	0.57
	FAM	313	42.38	4.53	0.26	0.78	0.74	0.82
	DEF	313	18.07	4.09	0.23	0.69	0.63	0.74
	GM	313	19.28	2.76	0.16	0.24	0.11	0.37
	SS	313	14.15	4.73	0.27	0.89	0.87	0.91
	TM	313	19.29	4.23	0.24	0.77	0.73	0.81
	REN	313	30.29	3.04	0.17	0.70	0.65	0.75
	SOC	313	26.05	3.65	0.21	0.63	0.57	0.70
	DIS	313	13.78	2.90	0.16	0.50	0.41	0.59
	HAR	313	21.08	2.39	0.14	0.08	-0.08	0.24
	TE	313	26.34	3.71	0.21	0.68	0.62	0.73
BFI	Conscientiousness	129	51.29	7.18	0.63	0.92	0.90	0.93
	Agreeableness	129	50.05	5.60	0.49	0.80	0.76	0.83
	Neuroticism	129	24.50	6.87	0.61	0.87	0.85	0.89
	Openness	129	47.78	7.91	0.70	0.91	0.90	0.92
	Extraversion	129	43.88	8.37	0.74	0.90	0.89	0.92
OCB	Total score	184	102.35	19.62	1.45	0.91	0.90	0.93
	Organizational	184	18.83	4.01	0.30	-	-	-
	People-directed	184	19.04	4.06	0.30	-	-	-
CWB	Total score	184	115.47	28.63	2.11	0.96	0.95	0.97
	Organizational	184	31.18	8.07	0.59	-	-	-
	People-directed	184	26.55	6.70	0.49	-	-	-
DASS-21	Depression	184	2.78	2.85	0.21	0.81	0.78	0.84
	Anxiety	184	4.49	3.29	0.24	0.73	0.68	0.77
	Stress	184	5.67	3.72	0.27	0.83	0.80	0.86
CPAI-2	FAC	313	6.55	1.95	0.11	-	-	-
	FAM	313	8.77	1.79	0.10	-	-	-
	DEF	313	2.31	2.04	0.12	-	-	-
	GM	313	8.01	2.16	0.12	-	-	-
	SS	313	4.22	1.98	0.11	-	-	-
	TM	313	6.67	2.83	0.16	-	-	-
	REN	313	9.61	1.67	0.09	-	-	-
	SOC	313	9.61	1.83	0.10	-	-	-
	DIS	313	6.96	1.79	0.10	-	-	-
	HAR	313	12.26	1.68	0.10	-	-	-
	TE	313	6.55	1.92	0.11	-	-	-

Table 6. (Continued)

					USA			
New CPAI	FAC	605	21.22	4.02	0.16	0.50	0.44	0.56
	FAM	605	36.69	7.21	0.29	0.80	0.78	0.83
	DEF	605	20.60	3.55	0.14	0.34	0.25	0.42
	GM	605	19.33	3.42	0.14	0.41	0.43	0.48
	SS	605	19.60	4.92	0.20	0.67	0.63	0.71
	TM	605	18.73	5.05	0.21	0.76	0.73	0.78
	REN	605	28.52	3.81	0.15	0.67	0.63	0.71
	SOC	605	26.64	4.38	0.18	0.70	0.66	0.74
	DIS	605	14.38	3.33	0.14	0.49	0.43	0.56
	HAR	605	17.06	3.52	0.14	0.47	0.41	0.53
	TE	605	25.17	4.31	0.18	0.63	0.58	0.67
BFI	Conscientiousness	257	44.95	9.41	0.59	0.90	0.89	0.91
	Agreeableness	257	45.98	7.70	0.48	0.82	0.80	0.84
	Neuroticism	257	33.21	11.07	0.69	0.92	0.91	0.93
	Openness	257	47.86	8.19	0.51	0.87	0.85	0.88
	Extraversion	257	39.19	9.99	0.62	0.88	0.87	0.90
OCB	Total score	345	95.91	25.19	1.36	0.94	0.93	0.95
	Organizational	345	17.41	5.08	0.27	-	-	-
	People-directed	345	18.52	5.00	0.27	-	-	-
CWB	Total score	108	119.85	38.11	3.67	0.96	0.95	0.96
	Organizational	108	33.61	10.91	1.05	-	-	-
	People-directed	108	26.31	9.10	0.88	-	-	-
DASS-21	Depression	345	12.85	5.45	0.29	0.93	0.92	0.94
	Anxiety	345	11.77	4.64	0.25	0.88	0.86	0.89
	Stress	345	13.56	4.80	0.26	0.87	0.85	0.89
CPAI-2	FAC	605	5.21	2.36	0.10	-	-	-
	FAM	605	5.46	2.47	0.10	-	-	-
	DEF	605	2.60	2.20	0.09	-	-	-
	GM	605	6.00	2.34	0.10	-	-	-
	SS	605	5.97	1.99	0.08	-	-	-
	TM	605	4.06	2.80	0.11	-	-	-
	REN	605	8.02	1.95	0.08	-	-	-
	SOC	605	8.06	1.88	0.08	-	-	-
	DIS	605	5.72	2.40	0.10	-	-	-
	HAR	605	10.29	2.16	0.09	-	-	-
	TE	605	6.02	1.64	0.07	-	-	-

Table 7. New CPAI item characteristics for all items in the full sample.

Scale	Name	Item	Overall Sample												
			N	M	SD	SE	M. r_{xx}	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
FAC							0.66	0.46							
	FAC1	It really bothers me when someone says something negative about me.	918	3.72	1.11	0.04			0.33	0.62	1.57	-2.62	-1.36	-0.68	1.00
	FAC2	I enjoy dressing nicely when I walk around in public.	918	3.95	1.04	0.03			0.47	0.38	0.21	-16.81	-10.05	-4.99	3.10
	FAC3	I buy things I can't afford to enhance my public image.	918	1.99	1.13	0.04			0.45	0.44	0.13	-1.82	8.26	13.84	26.61
	FAC4r	I do not feel embarrassed when I make a mistake in a group	918	3.51	1.24	0.04			0.44	0.44	0.81	-3.49	-1.45	-0.67	1.64
	FAC5	I rarely apologize even if I am wrong.	918	1.73	0.98	0.03			0.44	0.43	0.32	0.23	5.62	7.51	12.07
	FAC6	I often refuse help even if I need it.	918	2.77	1.24	0.04			0.45	0.43	0.60	-2.89	-0.08	1.16	4.20
	FAC7	I would be angry if someone pointed out my mistakes in front of others.	918	3.30	1.20	0.04			0.31	0.64	1.80	-1.97	-0.76	-0.04	1.38
FAM							0.84	0.82							
	FAM1	I often call my parents to talk about my daily life	918	3.46	1.35	0.05			0.80	0.65	1.51	-1.73	-0.83	-0.31	1.00
	FAM2	I feel a sense of duty to my family.	918	4.21	0.97	0.03			0.80	0.74	2.02	-2.66	-1.82	-1.23	0.17
	FAM3	When something important occurs in my life, I call my siblings or parents.	918	3.97	1.17	0.04			0.79	0.73	1.78	-2.21	-1.43	-0.95	0.41
	FAM4	When my family needs my support, I will do everything I can to help.	918	4.45	0.76	0.03			0.80	0.71	1.98	-3.33	-2.54	-1.79	-0.16
	FAM5	My family is really important to me.	918	4.53	0.83	0.03			0.80	0.76	2.62	-2.70	-2.05	-1.59	-0.45

Table 7. (Continued)

Scale	Name	Item	Overall Sample												
			N	M	SD	SE	M. r_{xx}	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	FAM6r	I have spoken ill of my family when talking to others.	918	3.50	1.33	0.04			0.81	0.60	1.19	-2.48	-0.89	-0.24	0.92
	FAM7	I take my family's reputation into account when deciding how to behave.	918	3.54	1.28	0.04			0.81	0.63	1.32	-2.01	-1.17	-0.55	1.19
	FAM8r	I would be irritated if I had to listen to a family member tell me about their problems.	918	3.98	1.01	0.03			0.83	0.45	0.85	-4.77	-2.66	-1.56	0.89
	FAM9	I do things to meet my parents' expectations regardless of my own preferences.	918	2.92	1.21	0.04			0.83	0.42	0.61	-3.10	-0.67	0.98	4.06
	FAM10r	I do not need my family's support.	918	4.08	1.15	0.04			0.80	0.69	1.68	-2.50	-1.61	-0.99	0.11
DEF							0.71	0.48							
	DEF1	When my friends and I end up in trouble it is usually their fault.	918	2.53	1.09	0.04			0.39	0.60	1.44	-1.32	0.07	1.33	2.82
	DEF2r	I blame myself for things that I am not responsible for.	918	3.05	1.27	0.04			0.63	0.10	-0.48	4.39	0.84	-0.72	-3.79
	DEF3	If I make a mistake, it's okay because others have probably made it too.	918	3.09	1.21	0.04			0.43	0.51	0.69	-3.25	-0.89	0.42	3.20
	DEF4	If I am not able to reach a deadline, it is typically because outside factors took up too much of my time.	918	3.01	1.22	0.04			0.38	0.61	1.29	-1.94	-0.48	0.37	2.05
	DEF5	If everything were fair, I would be more successful than I am now.	918	3.27	1.15	0.04			0.45	0.49	0.61	-4.11	-2.00	0.21	3.19

Table 7. (Continued)

Scale	Name	Item	Overall Sample												
			N	M	SD	SE	M. r_{xx}	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	DEF6	Often other people are the reason I can't succeed.	918	1.98	0.96	0.03			0.40	0.59	1.15	-0.65	1.18	2.43	4.08
	DEF7	When people criticize my work, it's usually because they don't understand it.	918	2.81	1.02	0.03			0.37	0.64	1.48	-1.95	-0.37	0.88	2.92
GM		If I have a problem with someone, I prefer to keep it to myself and let it go quickly.					0.66	0.32							
	GM1		918	3.02	1.23	0.04			0.37	0.38	-0.19	10.18	2.44	-1.50	-11.70
	GM2	If I saw someone cheating off my work, I would let them continue.	918	2.22	1.14	0.04			0.44	0.26	-0.53	1.48	-1.33	-3.07	-6.76
	GM3r	I doubt other people's intentions.	918	2.98	1.18	0.04			0.24	0.51	1.76	-1.74	-0.38	0.44	1.78
	GM4r	I hold grudges.	918	3.32	1.26	0.04			0.11	0.63	1.52	-2.17	-0.70	0.03	1.20
	GM5r	I cannot work with people if they make simple mistakes.	918	3.62	1.09	0.04			0.23	0.53	0.81	-4.28	-2.01	-0.71	1.82
	GM6	I believe people deserve second chances.	918	4.16	0.79	0.03			0.22	0.58	0.58	-8.57	-5.51	-3.32	1.20
SS	SS1r	I enjoy playing team sports.	918	2.31	1.33	0.04	0.83	0.78	0.76	0.61	1.39	-0.68	0.73	1.24	1.92
	SS2	I think collaborating with others can be a waste of time.	918	2.03	1.00	0.03			0.75	0.65	1.45	-0.71	1.09	1.84	3.25
	SS3	I do not like working on team projects.	918	2.49	1.30	0.04			0.72	0.74	2.70	-0.72	0.28	0.75	1.56
	SS4r	I prefer working with others rather than working independently.	918	2.89	1.25	0.04			0.74	0.66	1.82	-1.60	-0.15	0.52	1.55

Table 7. (Continued)

Scale	Name	Item	Overall Sample												
			N	M	SD	SE	M. r_{xx}	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	SS5r	I feel the same level of pride when my team achieves success as I would if I did it myself.	918	1.85	0.97	0.03			0.75	0.66	1.32	-0.30	1.47	2.21	3.45
	SS6	I am always worried that other teammates may take credit for my effort.	918	2.38	1.11	0.04			0.77	0.57	0.97	-1.51	0.64	1.63	3.79
	SS7r	I can share credit with people even if I did more work.	918	1.97	0.84	0.03			0.77	0.58	0.99	-1.18	1.82	3.02	4.60
	SS8r	I am happy when my team succeeds, even if I'm not personally recognized.	918	1.83	0.84	0.03			0.77	0.56	0.78	-0.75	2.44	3.82	6.51
TM		When I start a family, I					0.81	0.75							
	TM1r	want to make my own traditions rather then follow my family's.	918	2.43	1.07	0.04			0.73	0.61	1.10	-1.54	0.33	1.67	3.45
	TM2r	I am open to new ideas that are inconsistent with traditions.	918	1.87	0.88	0.03			0.75	0.53	0.87	-0.71	2.05	3.50	5.16
	TM3	I base my principles firmly in line with the principles of my parents.	918	2.98	1.16	0.04			0.70	0.72	1.96	-1.56	-0.49	0.44	1.83
	TM4	I think it is ignorant for people to doubt and challenge old ways.	918	2.14	1.08	0.04			0.74	0.58	1.15	-0.81	0.96	1.90	3.39
	TM5r	I think my parents' way of dealing with things are no longer appropriate.	918	3.04	1.19	0.04			0.73	0.60	1.19	-2.14	-0.66	0.47	2.09
	TM6	Traditions should always be followed.	918	2.74	1.12	0.04			0.69	0.74	2.18	-1.25	-0.25	0.74	2.15
	TM7	Teaching traditions to our children is important.	918	3.71	1.01	0.03			0.72	0.65	1.50	-3.15	-1.65	-0.51	1.11

REN	REN1	I will do my best to return someone's favor.	918	4.41	0.69	0.02	0.75	0.69	0.62	0.72	2.04	-3.58	-2.67	-1.94	0.04
	REN2	I will not forget even the tiniest help I received from someone else.	918	4.19	0.93	0.03			0.65	0.61	1.29	-3.77	-2.39	-1.60	0.31
	REN3	If a family member needed help getting a job, I would try to get them a job in my company.	918	3.80	1.05	0.04			0.72	0.41	0.49	-6.93	-3.92	-1.87	2.17
	REN4	If a friend gives me a gift, I have to give them a gift back at some point.	918	4.11	0.91	0.03			0.64	0.63	1.33	-3.77	-2.38	-1.46	0.54
	REN5	When a co-worker compliments me, I feel the need to compliment them as well.	918	3.93	1.00	0.03			0.65	0.60	1.16	-3.53	-2.23	-1.20	0.93
	REN6	I feel obligated to help people who helped me.	918	4.37	0.77	0.03			0.61	0.73	2.25	-3.24	-2.25	-1.62	0.03
	REN7	Maintaining long-term relationships are more important than getting instant rewards.	918	4.32	0.80	0.03			0.67	0.53	0.81	-6.15	-4.38	-2.62	0.16
SOC	SOC1	I am careful of how I act in public to avoid offending anyone.	918	3.99	0.95	0.03	0.77	0.68	0.68	0.47	0.63	-6.12	-3.93	-2.25	1.41
	SOC2	I am sensitive to other people's feelings and emotions.	918	4.03	0.95	0.03			0.61	0.69	1.39	-3.60	-2.16	-1.17	0.62
	SOC3	I know immediately if I say anything inappropriate in front of others.	918	4.04	0.90	0.03			0.65	0.58	0.70	-6.37	-3.75	-2.21	1.16
	SOC4r	It is difficult for me to interpret other people's feelings.	918	3.79	1.08	0.04			0.68	0.51	0.60	-5.72	-2.91	-1.68	1.82

Table 7. (Continued)

Scale	Name	Item	Overall Sample											
			N	M	SD	SE	M. r_{xx}	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	SOC5	If a coworker seemed upset, I would ask them if they wanted to talk about it.	918	3.95	0.95	0.03			0.64	0.59	1.02	-4.13	-2.48	-1.49	1.08
	SOC6	The emotions of people around me always influence my own emotions easily.	918	3.33	1.11	0.04			0.63	0.61	1.88	-2.11	-0.86	-0.05	1.55
	SOC7	If someone around me is upset, I would feel upset too.	918	3.31	1.12	0.04			0.61	0.66	2.31	-1.88	-0.78	-0.03	1.41
DIS	DIS1	It bothers me if my plans change at the last minute.	918	3.68	1.17	0.04	0.70	0.49	0.39	0.63	0.84	-3.74	-1.82	-0.97	1.42
	DIS2	I spend my day with a planned routine.	918	3.16	1.19	0.04			0.40	0.58	0.19	-12.60	-3.67	1.17	10.03
	DIS3r	I am flexible when unexpected things happen.	918	2.21	0.97	0.03			0.42	0.63	2.58	-0.93	0.67	1.38	2.43
	DIS4	I wake up in the mornings at the same time everyday.	918	3.15	1.29	0.04			0.52	0.47	-0.06	31.62	12.24	0.95	-35.48
	DIS5r	I can improvise well if I have to.	918	1.98	0.91	0.03			0.45	0.58	1.68	-0.73	1.23	2.04	3.01
HAR	HAR1	Stressful events don't usually bother me.	918	2.71	1.31	0.04	0.68	0.52	0.46	0.57	0.70	-2.06	0.11	0.97	3.53
	HAR2	I let things go to avoid interpersonal conflict.	918	3.62	1.08	0.04			0.43	0.62	1.27	-3.01	-1.49	-0.59	1.43
	HAR3r	I would help my friend fend off the attackers if they ever got into a fight.	918	1.88	0.93	0.03			0.55	0.40	0.38	-0.96	3.48	6.99	11.88
	HAR4	I seldom argue with people.	918	3.75	1.17	0.04			0.36	0.70	1.98	-2.24	-1.22	-0.55	0.64
	HAR5	If a coworker did something wrong, I wouldn't say anything to avoid upsetting them.	918	3.13	1.14	0.04			0.45	0.59	1.26	-2.53	-0.63	0.22	2.17

Table 7. (Continued)

Scale	Name	Item	Overall Sample												
			N	M	SD	SE	M. r_{xx}	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	HAR6r	I get upset with other people easily	918	3.33	1.16	0.04			0.58	0.36	0.16	-17.28	-5.98	0.29	9.65
TE							0.74	0.64							
	TE1	I find it important to save a little of my paycheck every month.	918	4.34	0.86	0.03			0.61	0.53	0.83	-5.48	-3.88	-2.64	-0.07
	TE2r	I always want to have the newest version of everything.	918	3.29	1.23	0.04			0.64	0.46	0.50	-5.12	-1.62	-0.03	3.10
	TE3	I don't spend money on things unless it is necessary.	918	3.36	1.24	0.04			0.55	0.66	1.69	-2.13	-0.72	-0.13	1.21
	TE4	I hold on to things just in case I need them someday.	918	3.96	0.96	0.03			0.65	0.42	0.48	-7.70	-4.69	-3.01	2.02
	TE5r	Spending money and buying things makes me feel happy.	918	2.52	1.15	0.04			0.58	0.59	0.95	-1.81	0.34	1.57	3.15
	TE6	I think carefully before I buy something.	918	4.03	1.00	0.03			0.54	0.71	2.39	-2.51	-1.53	-0.95	0.45
	TE7	I will try to fix something before I replace it.	918	4.07	0.89	0.03			0.60	0.57	0.90	-5.28	-3.07	-1.84	0.91

Note. M. r_{xx} = marginal reliability, Scale α = alpha reliability, α drop = alpha if item dropped, r. cor. = item-scale correlation, a = IRT discrimination parameter, b1 – b4 = IRT intercept parameters.

Table 8. Correlations among the new CPAI scales and the equivalent CPAI-2 emic scales.

Scale	r	[95% CI]
FAC	0.48	[.43, .53]
FAM	0.68	[.64, .71]
DEF	0.38	[.32, .43]
GM	0.46	[.40, .51]
SS	0.46	[.40, .51]
TM	0.54	[.49, .58]
REN	0.37	[.31, .42]
SOC	0.46	[.41, .51]
DIS	0.25	[.19, .31]
HAR	0.45	[.39, .50]
TE	0.52	[.47, .57]

Table 9. Correlations between BFI factors and the new CPAI and CPAI-2 scales.

Scale	Inventory	Agreeableness	Conscientiousness	Extraversion	Neuroticism	Openness
FAC						
	CPAI-2	-0.06	-0.1	-0.2	0.28	-0.18
	new CPAI	-0.26	-0.25	-0.24	0.46	-0.19
FAM						
	CPAI-2	0.48	0.53	0.37	-0.53	0.17
	new CPAI	0.48	0.5	0.36	-0.35	0.21
DEF						
	CPAI-2	-0.37	-0.17	-0.1	0.2	-0.26
	new CPAI	-0.32	-0.19	-0.08	0.04	-0.25
GM						
	CPAI-2	0.59	0.41	0.31	-0.52	0.27
	new CPAI	0.52	0.21	0.12	-0.26	0.22
SS						
	CPAI-2	-0.3	-0.24	-0.32	0.28	-0.04
	new CPAI	-0.54	-0.47	-0.47	0.49	-0.39
TM						
	CPAI-2	0.11	0.29	0.18	-0.3	-0.02
	new CPAI	0.17	0.31	0.15	-0.24	-0.05
REN						
	CPAI-2	0.28	0.27	0.04	-0.1	0.05
	new CPAI	0.48	0.38	0.25	-0.18	0.43
SOC						
	CPAI-2	0.58	0.46	0.45	-0.3	0.38
	new CPAI	0.45	0.22	0.13	0.04	0.38
DIS						
	CPAI-2	0.15	0.43	0.1	-0.23	-0.01
	new CPAI	-0.11	0.02	-0.27	0.27	-0.18
HAR						
	CPAI-2	0.6	0.56	0.22	-0.47	0.28
	new CPAI	0.41	0.37	0.1	-0.51	0.06
TE						
	CPAI-2	0.17	0.19	0	-0.13	0.12
	new CPAI	0.27	0.32	-0.02	-0.22	0.2

Table 10. Correlations between criterion scales and the new CPAI and CPAI-2 scales.

Factor	Inventory	CWB-O	CWB-P	CWB Total	OCB-O	OCB-P	OCB-Total	DASS-Anxiety	DASS-Depression	DASS-Stress
FAC										
	CPAI-2	0.01	0.05	0.03	0.02	0.02	0.02	0	0.01	0.1
	new CPAI	0.13	0.12	0.13	-0.1	-0.06	-0.1	0.24	0.22	0.32
FAM										
	CPAI-2	-0.33	-0.19	-0.28	0.19	0.09	0.18	-0.55	-0.64	-0.59
	new CPAI	-0.49	-0.35	-0.44	0.25	0.2	0.25	-0.34	-0.45	-0.38
DEF										
	CPAI-2	0.31	0.29	0.32	-0.06	-0.09	-0.11	0.27	0.25	0.3
	new CPAI	0.28	0.22	0.26	-0.16	-0.13	-0.17	0.2	0.23	0.23
GM										
	CPAI-2	-0.35	-0.26	-0.32	0.2	0.13	0.2	-0.49	-0.52	-0.56
	new CPAI	-0.22	-0.14	-0.19	0.21	0.19	0.22	-0.15	-0.18	-0.26
SS										
	CPAI-2	0.25	0.14	0.21	-0.13	-0.09	-0.14	0.36	0.39	0.39
	new CPAI	0.37	0.24	0.32	-0.32	-0.25	-0.33	0.47	0.54	0.54
TM										
	CPAI-2	-0.02	0.12	0.05	0.12	0	0.07	-0.23	-0.29	-0.28
	new CPAI	-0.2	-0.09	-0.15	0.1	-0.01	0.06	-0.07	-0.12	-0.11
REN										
	CPAI-2	-0.27	-0.21	-0.25	0.18	0.1	0.16	-0.16	-0.2	-0.16
	new CPAI	-0.44	-0.41	-0.44	0.25	0.23	0.26	-0.16	-0.2	-0.18
SOC										
	CPAI-2	-0.42	-0.32	-0.39	0.34	0.34	0.37	-0.34	-0.43	-0.39
	new CPAI	-0.23	-0.24	-0.25	0.25	0.34	0.31	0.02	-0.02	0.02
DIS										
	CPAI-2	-0.25	-0.14	-0.21	0.15	0.03	0.1	-0.15	-0.21	-0.15
	new CPAI	-0.08	-0.04	-0.07	-0.04	-0.1	-0.09	0.11	0.1	0.16
HAR										
	CPAI-2	-0.37	-0.31	-0.36	0.24	0.17	0.24	-0.43	-0.52	-0.5
	new CPAI	-0.13	-0.01	-0.08	0.12	0	0.08	-0.48	-0.52	-0.55
TE										
	CPAI-2	-0.14	-0.13	-0.14	0.03	0.05	0.03	0	0	-0.02
	new CPAI	-0.28	-0.29	-0.3	0.01	0.01	0.01	-0.14	-0.1	-0.19

Table 11. Item characteristics for items selected using the IFA approach.

Scale	Name	Item	N	M	SD	SE	M. r_{xx}	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
FAC							0.67	0.45							
	FAC1	It really bothers me when someone says something negative about me.	918	3.72	1.11	0.04			0.19	0.65	2.396	-2.145	-1.139	-0.574	0.829
	FAC3	I buy things I can't afford to enhance my public image.	918	1.99	1.13	0.04			0.55	0.09	0.132	-1.769	8.060	13.497	25.933
	FAC4r	I do not feel embarrassed when I make a mistake in a group	918	3.51	1.24	0.04			0.39	0.38	0.943	-3.082	-1.298	-0.601	1.450
	FAC7	I would be angry if someone pointed out my mistakes in front of others.	918	3.30	1.20	0.04			0.32	0.48	1.206	-2.488	-0.941	-0.054	1.717
FAM							0.80	0.83							
	FAM1	I often call my parents to talk about my daily life	918	3.46	1.35	0.05			0.82	0.60	1.528	-1.721	-0.829	-0.310	0.989
	FAM2	I feel a sense of duty to my family.	918	4.21	0.97	0.03			0.80	0.67	1.832	-2.806	-1.919	-1.301	0.171
	FAM3	When something important occurs in my life, I call my siblings or parents.	918	3.97	1.17	0.04			0.78	0.73	1.937	-2.131	-1.391	-0.924	0.397
	FAM4	When my family needs my support, I will do everything I can to help.	918	4.45	0.76	0.03			0.81	0.67	1.890	-3.400	-2.603	-1.827	-0.165
	FAM5	My family is really important to me.	918	4.53	0.83	0.03			0.79	0.78	2.754	-2.671	-2.039	-1.588	-0.449
	FAM10r	I do not need my family's support.	918	4.08	1.15	0.04			0.80	0.62	1.625	-2.550	-1.650	-1.022	0.106
DEF							0.68	0.63							
	DEF1	When my friends and I end up in trouble it is usually their fault.	918	2.53	1.09	0.04			0.58	0.49	1.262	-1.423	0.076	1.446	3.081
	DEF4	If I am not able to reach a deadline, it is typically because outside factors took up too much of my time.	918	3.01	1.22	0.04			0.56	0.52	1.290	-1.933	-0.470	0.377	2.057

Table 11. (Continued)

Scale	Name	Item	N	M	SD	SE	M. rxx	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	DEF5	If everything were fair, I would be more successful than I am now.	918	3.27	1.15	0.04			0.63	0.36	0.722	-3.523	-1.713	0.192	2.740
	DEF6	Often other people are the reason I can't succeed.	918	1.98	0.96	0.03			0.55	0.55	1.301	-0.599	1.098	2.249	3.751
	DEF7	When people criticize my work, it's usually because they don't understand it.	918	2.81	1.02	0.03			0.55	0.54	1.414	-1.991	-0.380	0.907	3.008
GM	GM3r	I doubt other people's intentions.	918	2.98	1.18	0.04	0.65	0.58	0.46	0.54	1.448	-1.937	-0.415	0.483	1.967
	GM4r	I hold grudges.	918	3.32	1.26	0.04			0.41	0.60	1.828	-1.972	-0.645	0.023	1.094
	GM5r	I cannot work with people if they make simple mistakes.	918	3.62	1.09	0.04			0.54	0.42	0.822	-4.241	-1.991	-0.697	1.807
	GM6	I believe people deserve second chances.	918	4.16	0.79	0.03			0.57	0.34	0.653	-7.673	-4.945	-2.991	1.088
SS	SS2	I think collaborating with others can be a waste of time.	918	2.03	1.00	0.03	0.82	0.71	0.65	0.56	1.242	-0.778	1.187	2.031	3.611
	SS3	I do not like working on team projects.	918	2.49	1.30	0.04			0.57	0.75	3.960	-0.664	0.239	0.678	1.443
	SS4r	I prefer working with others rather than working independently.	918	2.89	1.25	0.04			0.63	0.63	2.119	-1.491	-0.156	0.475	1.464
	SS7r	I can share credit with people even if I did more work.	918	1.97	0.84	0.03			0.69	0.48	0.788	-1.401	2.179	3.637	5.585
	SS8r	I am happy when my team succeeds, even if I'm not personally recognized.	918	1.83	0.84	0.03			0.71	0.39	0.487	-1.108	3.714	5.861	10.111
TM	TM1r	When I start a family, I want to make my own traditions rather then follow my family's.	918	2.43	1.07	0.04	0.80	0.75	0.73	0.48	1.076	-1.555	0.337	1.692	3.501

Table 11. (Continued)

Scale	Name	Item	N	M	SD	SE	M. rxx	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	TM3	I base my principles firmly in line with the principles of my parents.	918	2.98	1.16	0.04			0.68	0.69	2.040	-1.530	-0.486	0.429	1.807
	TM4	I think it is ignorant for people to doubt and challenge old ways.	918	2.14	1.08	0.04			0.75	0.42	1.068	-0.855	1.010	2.001	3.598
	TM5r	I think my parents' way of dealing with things are no longer appropriate.	918	3.04	1.19	0.04			0.73	0.52	1.223	-2.096	-0.648	0.455	2.051
	TM6	Traditions should always be followed.	918	2.74	1.12	0.04			0.68	0.72	2.147	-1.257	-0.259	0.744	2.178
	TM7	Teaching traditions to our children is important.	918	3.71	1.01	0.03			0.71	0.59	1.539	-3.093	-1.621	-0.503	1.103
REN	REN1	I will do my best to return someone's favor.	918	4.41	0.69	0.02	0.74	0.72	0.66	0.65	1.980	-3.632	-2.711	-1.971	0.041
	REN2	I will not forget even the tiniest help I received from someone else.	918	4.19	0.93	0.03			0.69	0.53	1.299	-3.749	-2.382	-1.597	0.306
	REN4	If a friend gives me a gift, I have to give them a gift back at some point.	918	4.11	0.91	0.03			0.68	0.56	1.343	-3.734	-2.358	-1.452	0.535
	REN5	When a co-worker compliments me, I feel the need to compliment them as well.	918	3.93	1.00	0.03			0.70	0.48	1.152	-3.556	-2.244	-1.204	0.936
	REN6	I feel obligated to help people who helped me.	918	4.37	0.77	0.03			0.64	0.71	2.311	-3.202	-2.234	-1.612	0.032
	REN7	Maintaining long-term relationships are more important than getting instant rewards.	918	4.32	0.80	0.03			0.72	0.37	0.783	-6.336	-4.512	-2.699	0.161
SOC	SOC2	I am sensitive to other people's feelings and emotions.	918	4.03	0.95	0.03	0.79	0.69	0.66	0.49	1.053	-4.372	-2.608	-1.398	0.742

Table 11. (Continued)

Scale	Name	Item	N	M	SD	SE	M. rxx	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	SOC5	If a coworker seemed upset, I would ask them if they wanted to talk about it.	918	3.95	0.95	0.03			0.70	0.42	0.837	-4.850	-2.894	-1.736	1.257
	SOC6	The emotions of people around me always influence my own emotions easily.	918	3.33	1.11	0.04			0.58	0.68	2.216	-1.980	-0.815	-0.040	1.448
	SOC7	If someone around me is upset, I would feel upset too.	918	3.31	1.12	0.04			0.54	0.74	3.034	-1.737	-0.722	-0.022	1.299
DIS	DIS1	It bothers me if my plans change at the last minute.	918	3.68	1.17	0.04	0.70	0.59	0.66	0.39	0.810	-3.850	-1.869	-1.002	1.462
	DIS3r	I am flexible when unexpected things happen.	918	2.21	0.97	0.03			0.34	0.67	2.600	-0.931	0.674	1.376	2.423
	DIS5r	I can improvise well if I have to.	918	1.98	0.91	0.03			0.48	0.58	1.693	-0.723	1.229	2.032	2.996
HAR	HAR2	I let things go to avoid interpersonal conflict.	918	3.62	1.08	0.04	0.65	0.57	0.48	0.49	1.289	-2.978	-1.479	-0.594	1.407
	HAR3r	I would help my friend fend off the attackers if they ever got into a fight.	918	1.88	0.93	0.03			0.63	0.18	0.402	-0.915	3.328	6.679	11.337
	HAR4	I seldom argue with people.	918	3.75	1.17	0.04			0.39	0.61	1.792	-2.353	-1.274	-0.578	0.670
	HAR5	If a coworker did something wrong, I wouldn't say anything to avoid upsetting them.	918	3.13	1.14	0.04			0.42	0.57	1.370	-2.407	-0.597	0.210	2.058
TE	TE1	I find it important to save a little of my paycheck every month.	918	4.34	0.86	0.03	0.73	0.66	0.65	0.39	0.819	-5.527	-3.916	-2.664	-0.068
	TE3	I don't spend money on things unless it is necessary.	918	3.36	1.24	0.04			0.55	0.64	1.735	-2.100	-0.712	-0.135	1.195
	TE5r	Spending money and buying things makes me feel happy.	918	2.52	1.15	0.04			0.64	0.41	0.877	-1.928	0.361	1.676	3.354

Table 11. (Continued)

Scale	Name	Item	N	M	SD	SE	M. rxx	Scale α	α drop	r. cor.	a	b1	b2	b3	b4
	TE6	I think carefully before I buy something.	918	4.03	1.00	0.03			0.53	0.71	2.518	-2.465	-1.503	-0.934	0.442
	TE7	I will try to fix something before I replace it.	918	4.07	0.89	0.03			0.64	0.41	0.851	-5.553	-3.220	-1.928	0.949

Note. M. r_{xx} = marginal reliability, Scale α = alpha reliability, α drop = alpha if item dropped, r. cor. = correlation of item with scale, a = IRT discrimination parameter, b1 – b4 = IRT intercept parameters.

Table 12. Correlations among the IFA-shortened new CPAI scales and the equivalent CPAI-2 emic scales.

Scale	r	[95% CI]
FAC	0.53	[.48, .57]
FAM	0.64	[.60, .67]
DEF	0.44	[.39, .49]
GM	0.62	[.58, .66]
SS	0.46	[.40, .51]
TM	0.54	[.49, .58]
REN	0.37	[.32, .43]
SOC	0.35	[.29, .41]
DIS	0.09	[.03, .15]
HAR	0.40	[.34, .45]
TE	0.49	[.44, .54]

Table 13. Correlations between BFI factors and the IFA-shortened new CPAI scales.

Scale	Agreeableness	Conscientiousness	Extraversion	Neuroticism	Openness
FAC	-0.11	-0.19	-0.24	0.43	-0.17
FAM	0.43	0.46	0.34	-0.30	0.20
DEF	-0.36	-0.27	-0.19	0.24	-0.29
GM	0.64	0.42	0.33	-0.47	0.33
SS	-0.53	-0.42	-0.40	0.42	-0.37
TM	0.20	0.34	0.18	-0.27	-0.01
REN	0.47	0.41	0.22	-0.20	0.43
SOC	0.31	0.02	0.06	0.18	0.30
DIS	-0.31	-0.29	-0.43	0.47	-0.35
HAR	0.34	0.24	-0.09	-0.25	-0.01
TE	0.26	0.38	0.03	-0.29	0.18

Table 14. Correlations between criterion scales and the IFA-shortened new CPAI scales.

Factor	Inventory	CWB-O	CWB-P	CWB Total	OCB-O	OCB-P	OCB-Total	DASS-Anxiety	DASS-Depression	DASS-Stress
FAC	0.04	0.04	0.04	-0.08	-0.04	-0.08	0.15	0.13	0.24	0.04
FAM	-0.45	-0.35	-0.42	0.22	0.19	0.23	-0.29	-0.40	-0.31	-0.45
DEF	0.26	0.21	0.25	-0.14	-0.10	-0.15	0.31	0.32	0.36	0.26
GM	-0.31	-0.20	-0.27	0.30	0.24	0.31	-0.36	-0.41	-0.45	-0.31
SS	0.36	0.22	0.31	-0.34	-0.26	-0.35	0.43	0.51	0.50	0.36
TM	-0.21	-0.10	-0.17	0.12	0.02	0.09	-0.10	-0.16	-0.15	-0.21
REN	-0.44	-0.38	-0.43	0.22	0.22	0.24	-0.20	-0.24	-0.21	-0.44
SOC	-0.10	-0.10	-0.10	0.20	0.31	0.26	0.12	0.09	0.14	-0.10
DIS	0.03	-0.03	0.00	-0.17	-0.17	-0.19	0.26	0.28	0.35	0.03
HAR	-0.13	-0.04	-0.10	0.07	-0.01	0.04	-0.37	-0.41	-0.39	-0.13
TE	-0.29	-0.27	-0.29	0.02	-0.01	0.00	-0.18	-0.15	-0.23	-0.29

Table 15. Cross-validation correlations between GA-shortened scales and CPAI-2 scale scores at a series of item costs.

Item Cost	FAC	FAM	DEF	GM	SS	TM	REN	SOC	DIS	HAR	TE	Average
0.005	0.58	0.73	0.44	0.68	0.46	0.60	0.49	0.69	0.56	0.67	0.44	0.58
0.01	0.57	0.74	0.54	0.68	0.49	0.64	0.53	0.64	0.54	0.67	0.41	0.59
0.02	0.57	0.71	0.54	0.69	0.49	0.60	0.55	0.64	0.57	0.64	0.41	0.58
0.03	0.55	0.74	0.53	0.65	0.44	0.59	0.47	0.61	0.50	0.62	0.42	0.56
0.04	0.54	0.69	0.49	0.66	0.43	0.61	0.54	0.61	0.53	0.67	0.36	0.56
0.05	0.50	0.71	0.54	0.65	0.48	0.61	0.43	0.63	0.47	0.64	0.33	0.54
0.06	0.53	0.72	0.50	0.62	0.47	0.59	0.49	0.58	0.49	0.67	0.31	0.54
0.07	0.44	0.64	0.57	0.62	0.46	0.54	0.48	0.61	0.48	0.69	0.27	0.53
0.08	0.52	0.66	0.50	0.63	0.49	0.56	0.48	0.58	0.53	0.60	0.31	0.53
0.09	0.39	0.71	0.50	0.63	0.45	0.56	0.44	0.54	0.47	0.66	0.31	0.51

Table 16. Number of items in retained from each of the GA-shortened scales scales at resulting from a series of item costs.

Item Cost	FAC	FAM	DEF	GM	SS	TM	REN	SOC	DIS	HAR	TE	Total
0.005	4	4	1	2	2	4	1	2	1	1	5	27
0.01	3	3	1	3	1	5	1	1	0	2	4	24
0.02	3	3	2	1	2	4	2	1	1	0	3	22
0.03	3	2	2	1	1	3	1	1	1	2	3	20
0.04	3	2	2	1	1	4	1	0	1	1	2	18
0.05	3	1	2	1	1	3	1	1	0	1	2	16
0.06	2	2	2	1	1	3	1	0	1	1	1	15
0.07	2	1	2	1	1	3	3	1	1	1	0	16
0.08	2	2	2	1	1	2	0	2	1	1	1	15
0.09	2	3	2	1	1	1	0	1	1	1	1	14

Note. The retained solution is highlighted.

Table 17. Item characteristics for items selected using the GA approach.

Scale	Name	Item	N	M	SD	SE	CV *r*	M. r_{xx}	Scale α	α drop	*r* cor.	a	b1	b2	b3	b4
FAC	FAC1	It really bothers me when someone says something negative about me.	918	3.72	1.11	0.04	0.55	0.54	0.32	*-0.01*	0.51	1.782	-2.449	-1.282	-0.644	0.934
	FAC3	I buy things I can't afford to enhance my public image.	918	1.99	1.13	0.04				*0.47*	0.11	0.150	-1.558	7.118	11.914	22.875
	FAC4r	I do not feel embarrassed when I make a mistake in a group	918	3.51	1.24	0.04				*0.19*	0.40	1.108	-2.738	-1.161	-0.536	1.300
FAM	FAM4	When my family needs my support, I will do everything I can to help.	918	4.45	0.76	0.03	0.74	0.54	0.49	*0.22*	0.50	1.562	-3.698	-2.849	-2.028	-0.249
	FAM7	I take my family's reputation into account when deciding how to behave.	918	3.54	1.28	0.04				*0.63*	0.50	1.625	-1.802	-1.068	-0.516	1.043
DEF	DEF1	When my friends and I end up in trouble it is usually their fault.	918	2.53	1.09	0.04	0.53		0.11	*0.06*	0.18	0.442	-3.341	0.108	3.342	7.454
DEF	DEF5	If everything were fair, I would be more successful than I am now.	918	3.27	1.15	0.04				*0.06*	0.18	0.548	-4.490	-2.170	0.259	3.499
GM	GM4r	I hold grudges.	918	3.32	1.26	0.04	0.65			-	-	-	-	-	-	-
SS	SS3	I do not like working on team projects.	918	2.49	1.30	0.04	0.44			-	-	-	-	-	-	-
TM	TM3	I base my principles firmly in line with the principles of my parents.	918	2.98	1.16	0.04	0.59	0.74	0.68	*0.59*	0.60	1.606	-1.718	-0.543	0.482	2.019
	TM4	I think it is ignorant for people to doubt and challenge old ways.	918	2.14	1.08	0.04				*0.67*	0.51	1.304	-0.756	0.887	1.752	3.113
	TM6	Traditions should always be followed.	918	2.74	1.12	0.04				*0.47*	0.69	2.599	-1.186	-0.229	0.708	1.999

Table 17. (Continued)

Scale	Name	Item	N	M	SD	SE	CV r	M. rxx	Scale α	*α drop*	r cor.	a	b1	b2	b3	b4
REN	REN6	I feel obligated to help people who helped me.	918	4.37	0.77	0.03	0.47			-	-	-	-	-	-	-
SOC	SOC5	If a coworker seemed upset, I would ask them if they wanted to talk about it.	918	3.95	0.95	0.03	0.61			-	-	-	-	-	-	-
DIS	DIS2	I spend my day with a planned routine.	918	3.16	1.19	0.04	0.50			-	-	-	-	-	-	-
HAR	HAR1	Stressful events don't usually bother me.	918	2.71	1.31	0.04	0.62	0.42	0.41	*0.29*	0.40	1.135	-1.431	0.073	0.669	2.421
	HAR4	I seldom argue with people.	918	3.75	1.17	0.04				*0.23*	0.40	1.109	-3.133	-1.645	-0.752	0.860
TE	TE2r	I always want to have the newest version of everything.	918	3.29	1.23	0.04	0.42	0.67	0.50	*0.66*	0.21	0.375	-6.744	-2.112	-0.038	4.063
	TE3	I don't spend money on things unless it is necessary.	918	3.36	1.24	0.04				*0.25*	0.61	1.803	-2.067	-0.699	-0.131	1.168
	TE6	I think carefully before I buy something.	918	4.03	1.00	0.03				*0.23*	0.63	2.265	-2.564	-1.555	-0.975	0.449

Note. CV *r* = Cross-Validation correlation between the New CPAI item and the equivalent CPAI-2 scale; M. r_{xx} = marginal reliability; Scale α = alpha reliability of the scale; α drop = alpha of the scale if item dropped, *r* cor. = correlation of item with scale, a = IRT discrimination parameter, b1 – b4 = IRT intercept parameters.

Table 18. Correlations among the GA-shortened new CPAI scales and the equivalent CPAI-2 emic scales.

Scale	r	[95% CI]
FAC	0.54	[.49, .58]
FAM	0.53	[.49, .58]
DEF	0.36	[.30, .41]
GM	0.55	[.50, .59]
SS	0.47	[.42, .52]
TM	0.56	[.51, .60]
REN	0.30	[.24, .35]
SOC	0.37	[.32, .43]
DIS	0.32	[.26, .38]
HAR	0.43	[.37, .48]
TE	0.43	[.38, .48]

Table 19. Correlations between BFI factors and the GA-shortened new CPAI scales.

Scale	Agreeableness	Conscientiousness	Extraversion	Neuroticism	Openness
FAC	0.01	-0.10	-0.19	0.33	-0.12
FAM	0.42	0.39	0.25	-0.34	0.17
DEF	-0.29	-0.18	-0.19	0.19	-0.25
GM	0.47	0.35	0.30	-0.44	0.24
SS	-0.37	-0.32	-0.36	0.36	-0.25
TM	0.14	0.27	0.16	-0.23	-0.06
REN	0.30	0.23	0.04	-0.04	0.25
SOC	0.32	0.17	0.29	-0.10	0.40
DIS	0.11	0.34	0.08	-0.05	0.11
HAR	0.35	0.39	0.22	-0.60	0.10
TE	0.22	0.29	-0.02	-0.22	0.13

Table 20. Correlations between criterion scales and the GA-shortened new CPAI scales.

Factor	CWB-O	CWB-P	CWB Total	OCB-O	OCB-P	OCB-Total	DASS-Anxiety	DASS-Depression	DASS-Stress
FAC	-0.05	-0.02	-0.04	-0.01	0.01	-0.01	0.05	0.02	0.13
FAM	-0.35	-0.23	-0.31	0.26	0.18	0.25	-0.33	-0.42	-0.36
DEF	0.22	0.18	0.21	-0.13	-0.11	-0.15	0.25	0.27	0.29
GM	-0.25	-0.12	-0.20	0.28	0.21	0.28	-0.31	-0.32	-0.38
SS	0.29	0.14	0.23	-0.22	-0.12	-0.21	0.47	0.55	0.51
TM	-0.16	-0.01	-0.10	0.12	0.03	0.09	-0.07	-0.14	-0.11
REN	-0.26	-0.25	-0.27	0.16	0.19	0.19	-0.15	-0.18	-0.16
SOC	-0.21	-0.19	-0.21	0.31	0.35	0.35	-0.08	-0.14	-0.13
DIS	-0.13	-0.05	-0.10	0.09	0.04	0.07	0.02	-0.03	-0.01
HAR	-0.17	-0.03	-0.11	0.17	0.04	0.13	-0.53	-0.57	-0.59
TE	-0.21	-0.20	-0.21	-0.03	-0.03	-0.04	-0.13	-0.09	-0.17

Table 21. Items retained by the IFA and GA approaches.

Scale	IFA items	GA items
FAC	**1, 3, 4**, 7	**1, 3, 4**
FAM	1, 2, 3, **4**, 5, 10	**4**, 7
DEF	**1**, 4, **5**, 6, 7	**1, 5**
GM	3, **4**, 5, 6	**4**
SS	2, **3**, 4, 7, 8	**3**
TM	1, **3, 4**, 5, **6**, 7	**3, 4, 6**
REN	1, 2, 4, 5, **6**, 7	**6**
SOC	2, **5**, 6, 7	**5**
DIS	1, 3, 5	2
HAR	2, 3, **4**, 5	1, **4**
TE	1, **3**, 5, **6**, 7	2, **3, 6**

Note. Bolded numbers represent items that are shared across the two approaches.

Appendix B:
Figures

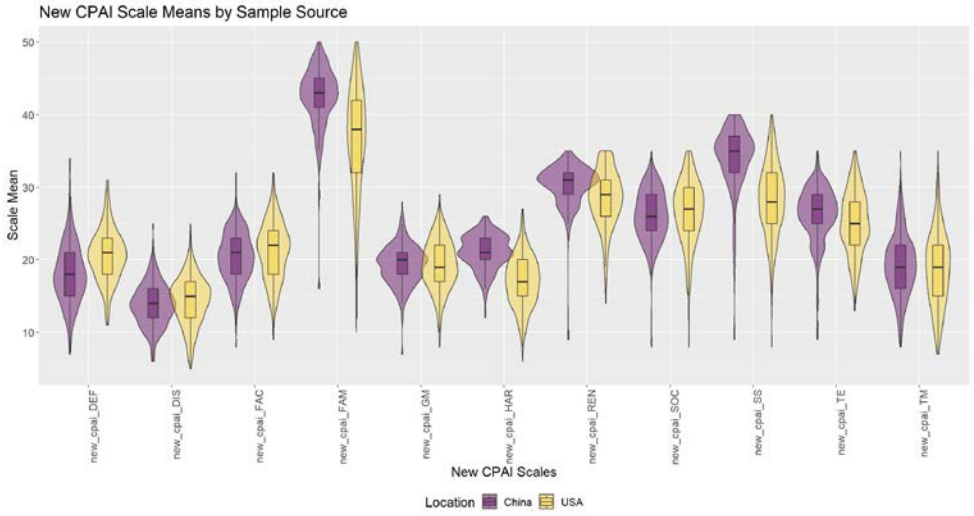

Figure 1. Distribution of New CPAI scale means among Chinese and US samples.

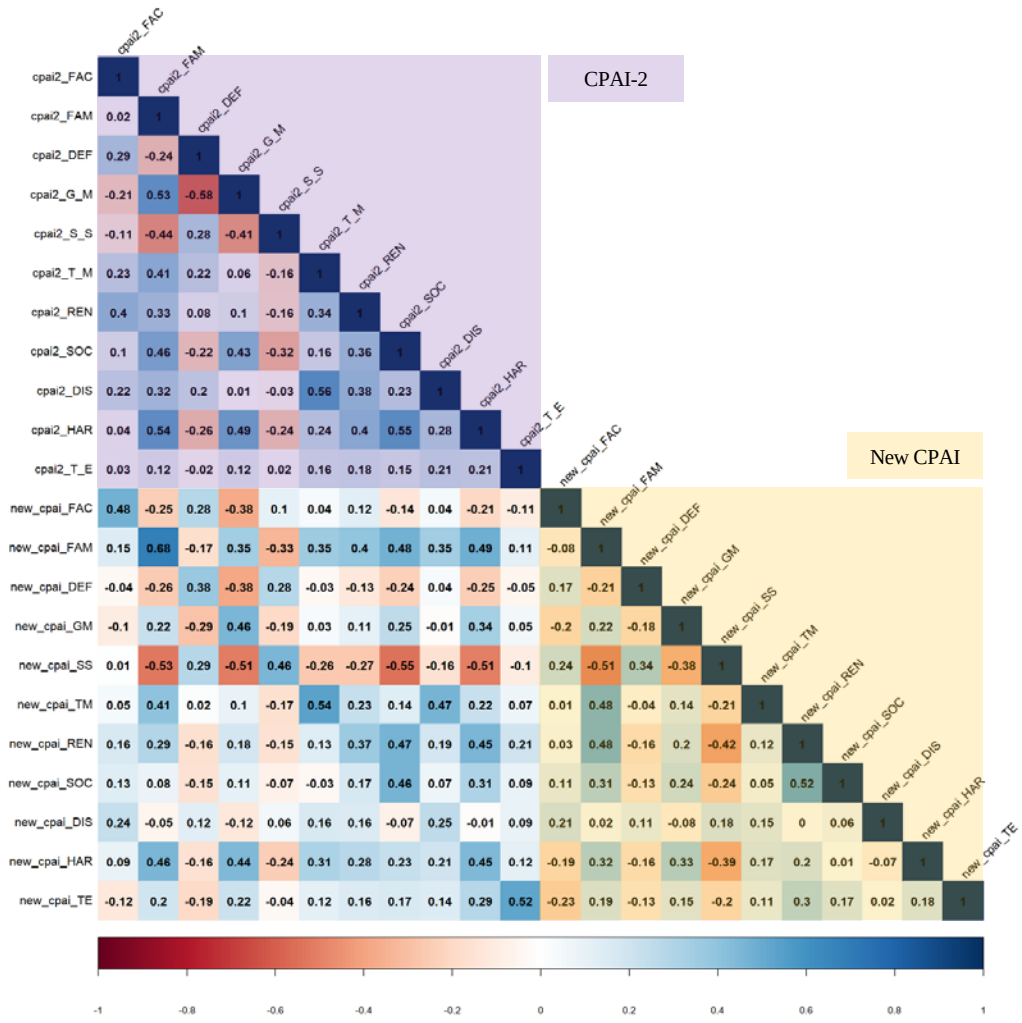

Figure 2. Correlations and intercorrelations among New CPAI and CPAI-2 scales.

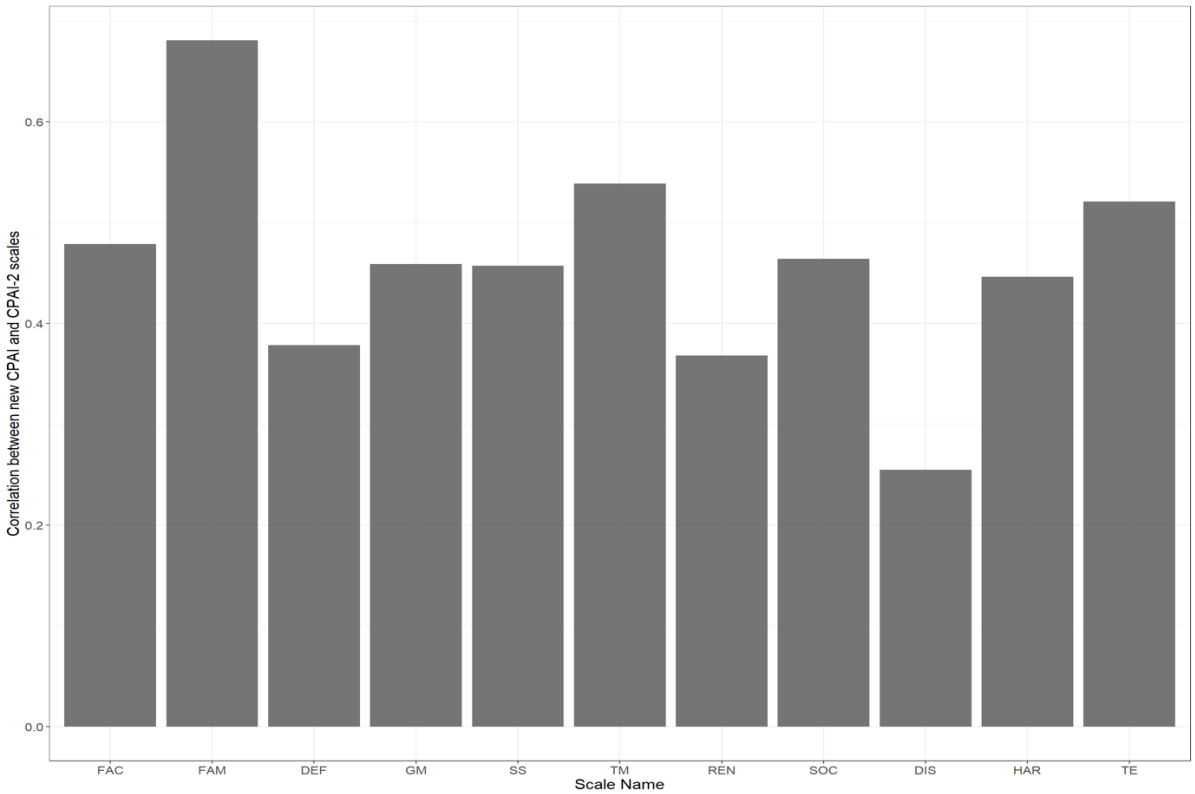

Figure 3. Correlations between New CPAI scales and equivalent CPAI-2 factors.

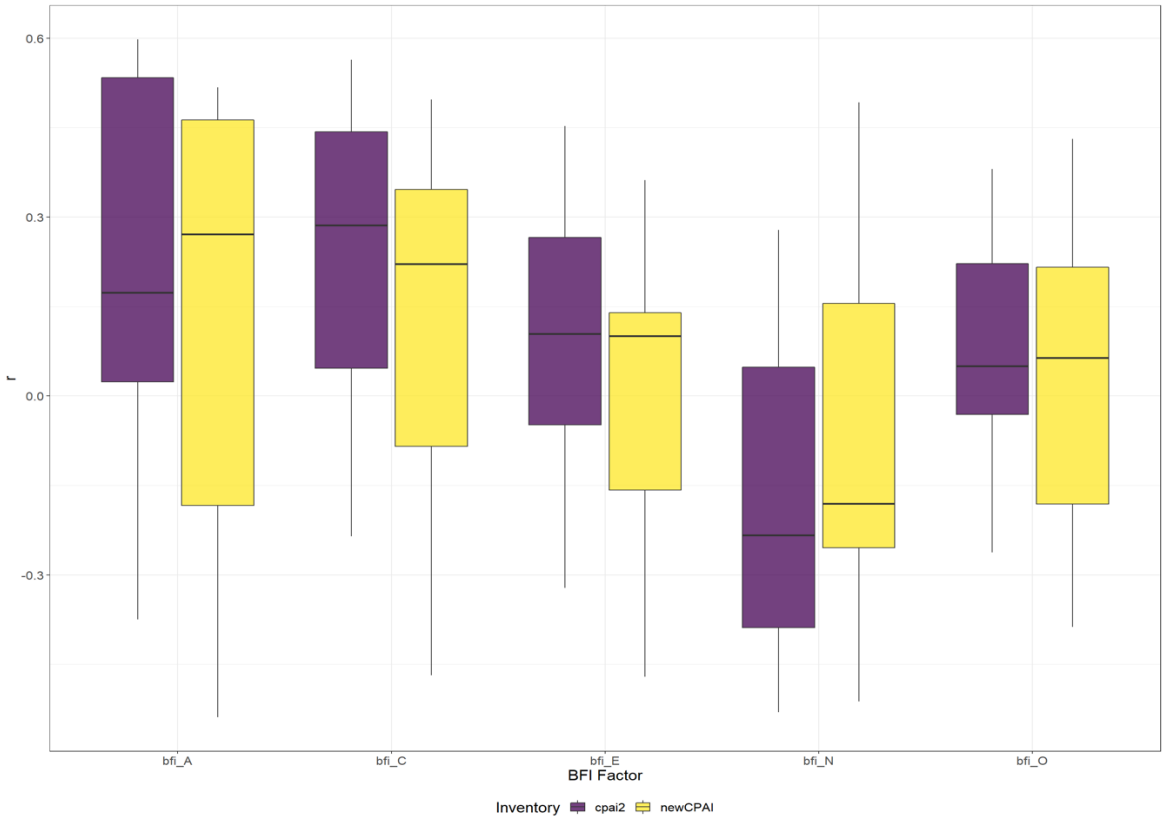

Figure 4. Distributions of correlations between BFI factors and New CPAI and CPAI-2 scales.

Figure 5. Correlations between BFI factors and all New CPAI and CPAI-2 scales.

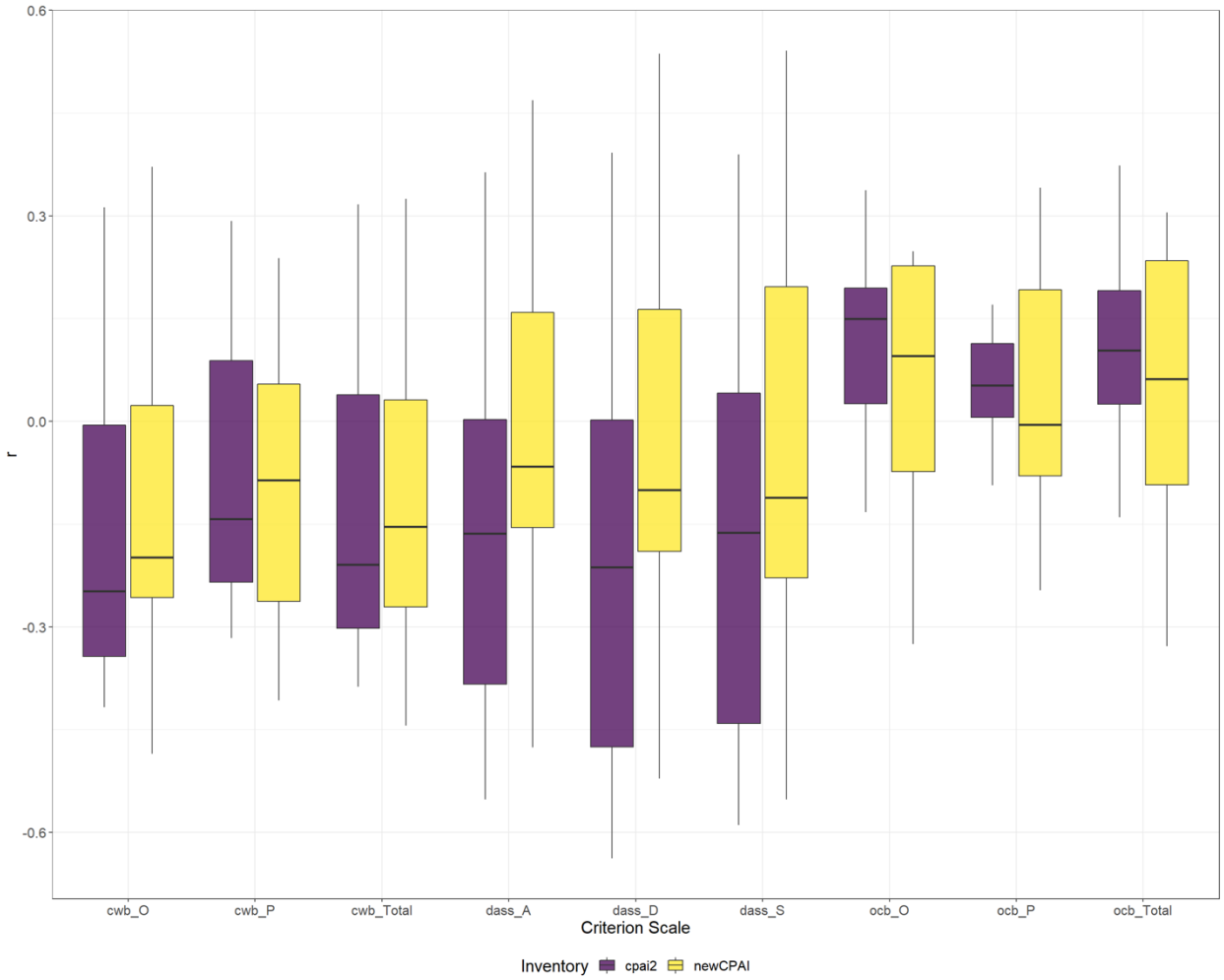

Figure 6. Distributions of correlations between criterion scales and New CPAI and CPAI-2 scales.

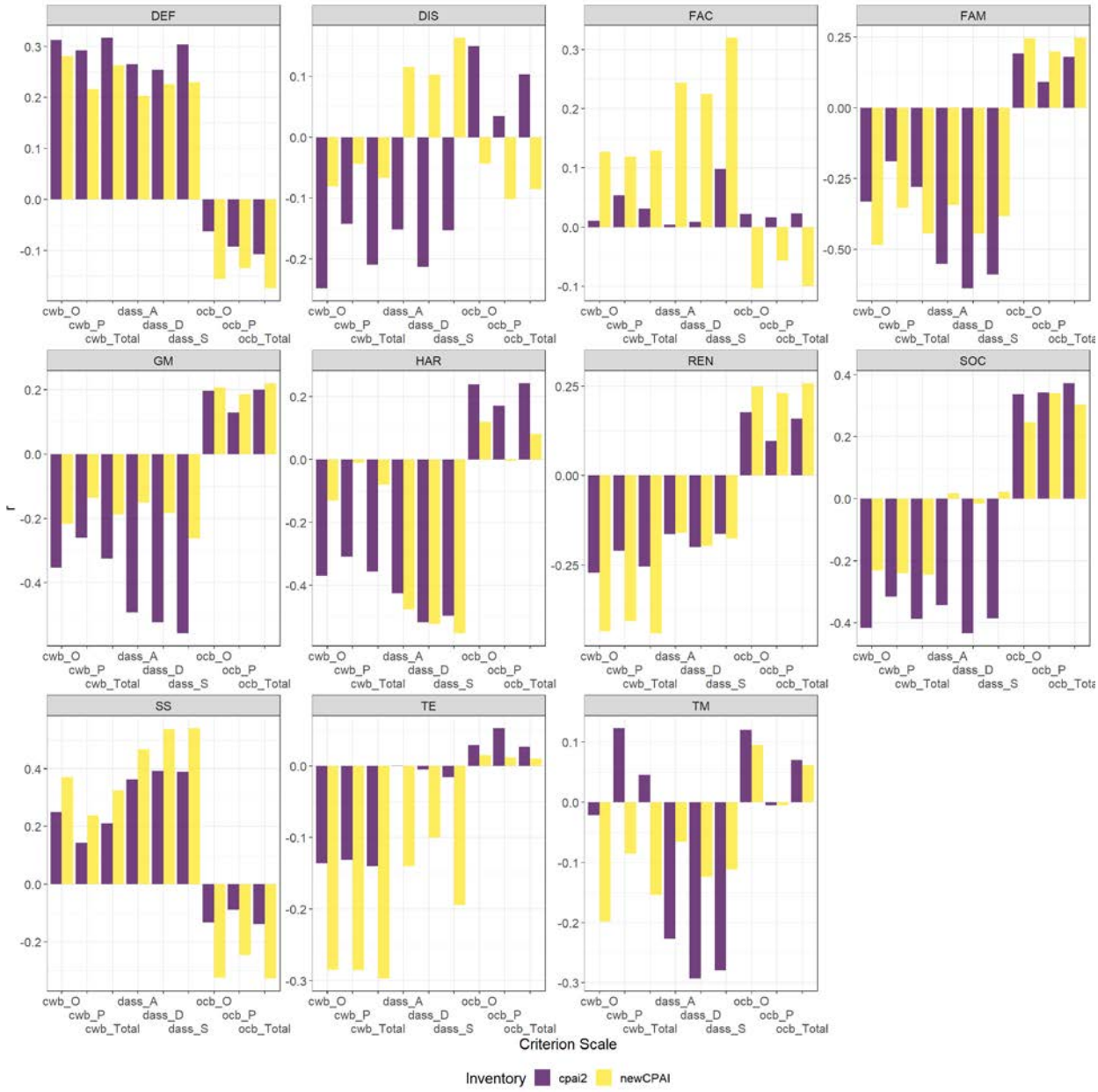

Figure 7. Correlations between criterion scales and all New CPAI and CPAI-2 scales.

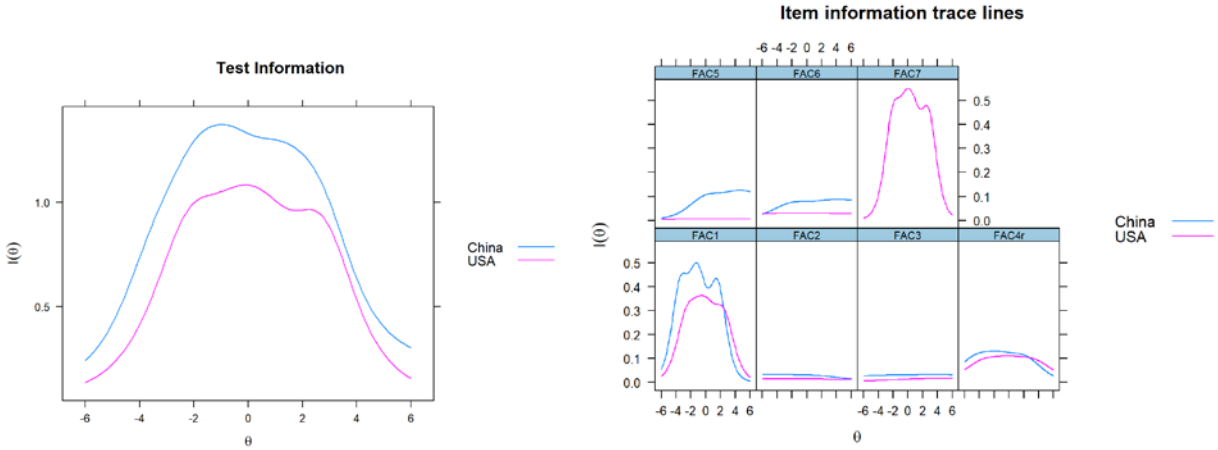

Figure 8. Scale information and item trace plots for the FAC scale.

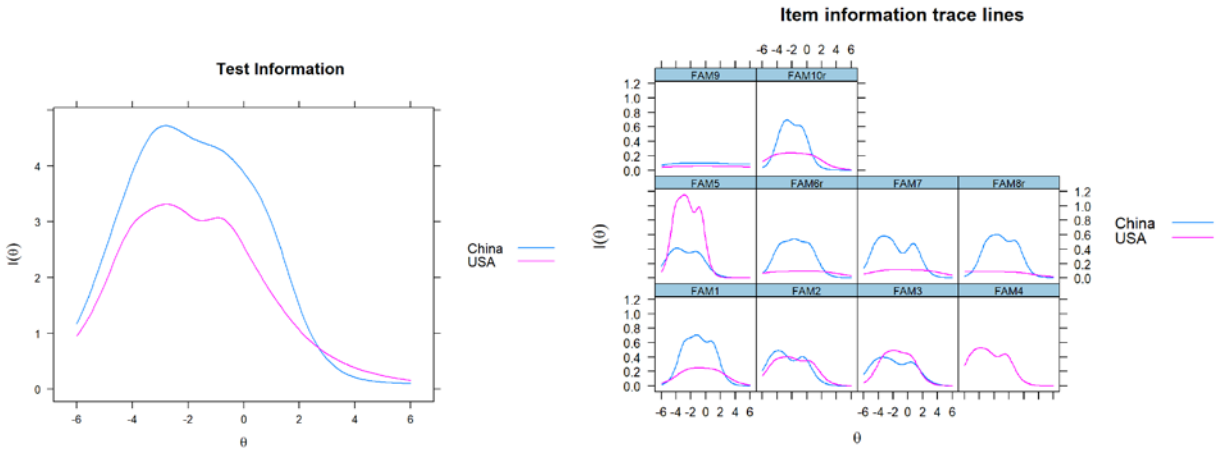

Figure 9. Scale information and item trace plots for the FAM scale.

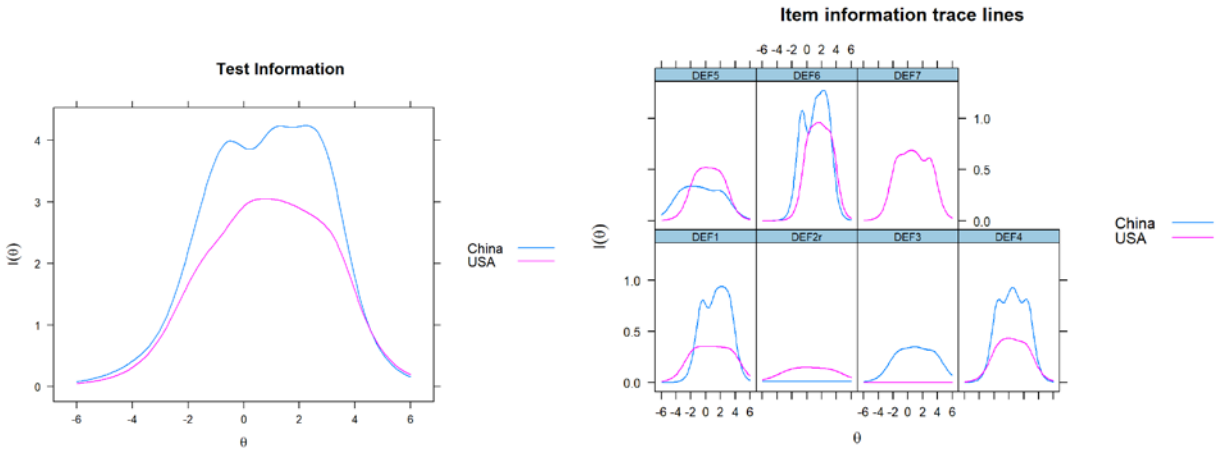

Figure 10. Scale information and item trace plots for the DEF scale.

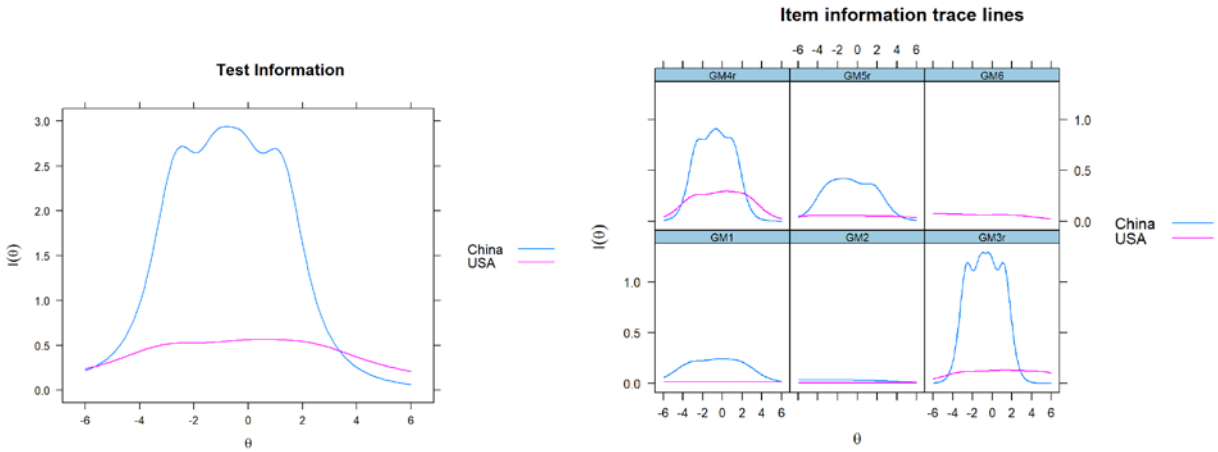

Figure 11. Scale information and item trace plots for the GM scale.

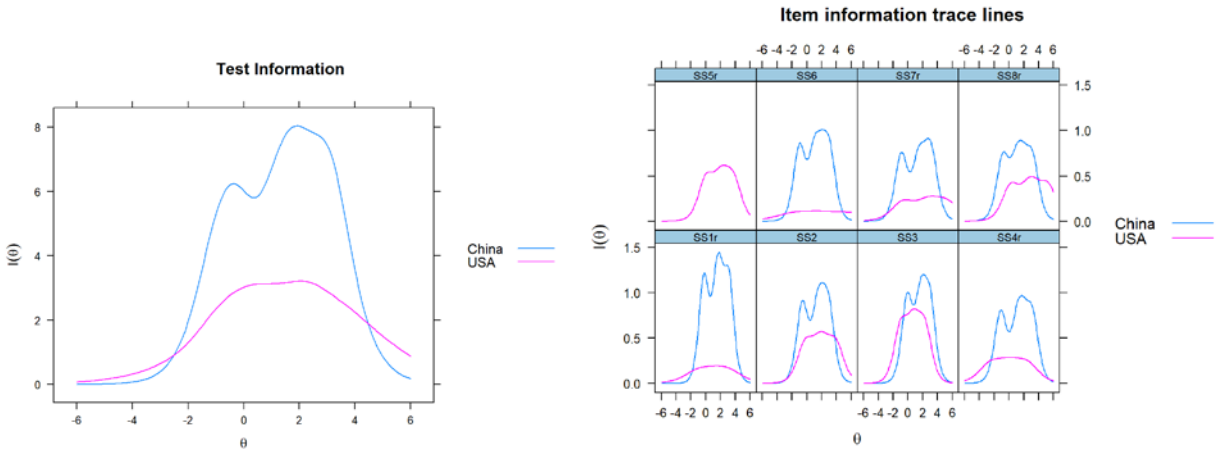

Figure 12. Scale information and item trace plots for the SS scale.

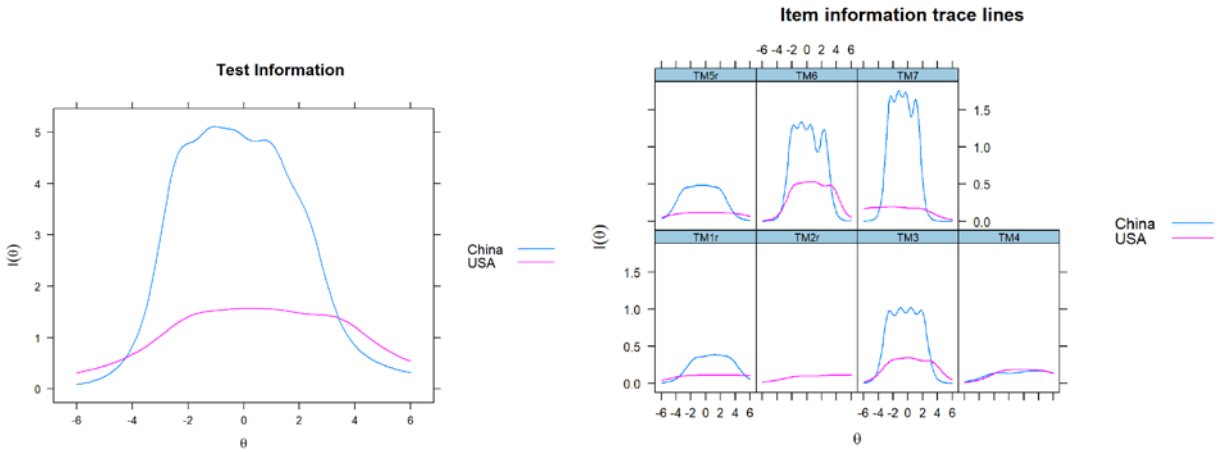

Figure 13. Scale information and item trace plots for the TM scale.

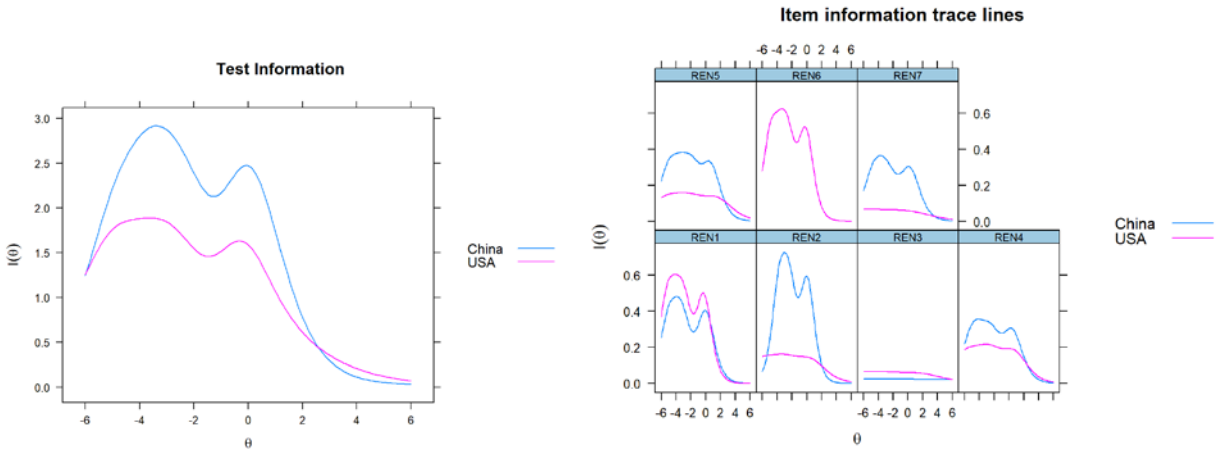

Figure 14. Scale information and item trace plots for the REN scale.

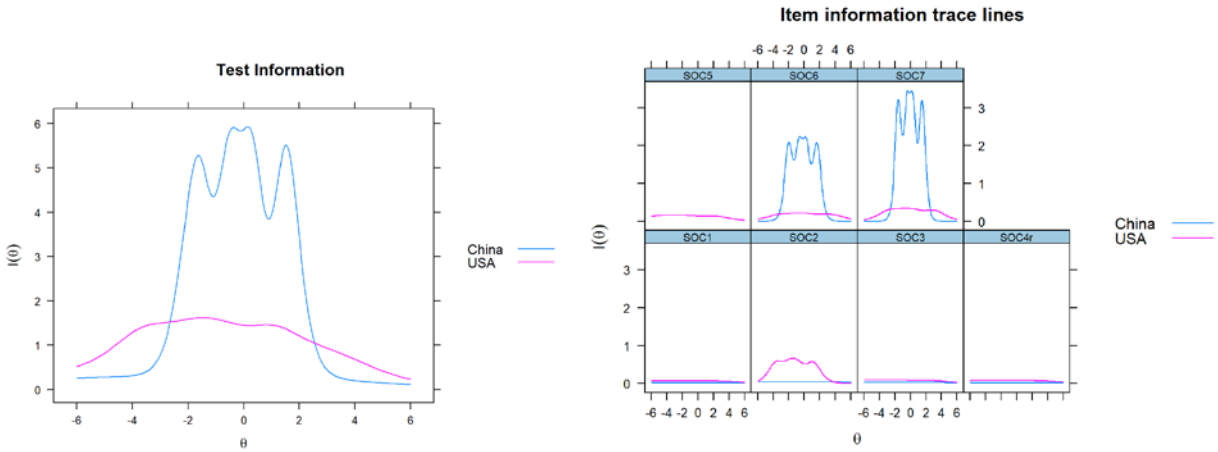

Figure 15. Scale information and item trace plots for the SOC scale.

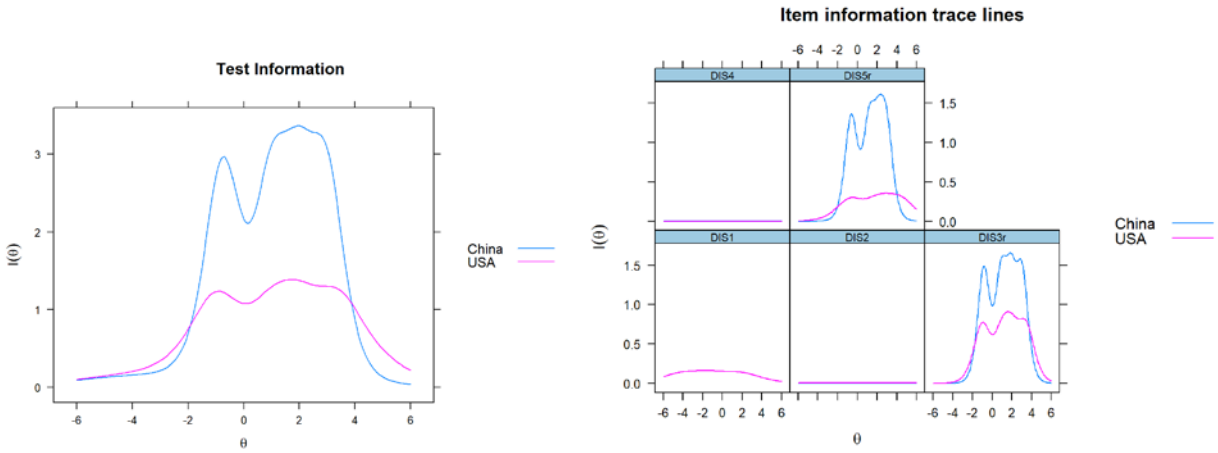

Figure 16. Scale information and item trace plots for the DIS scale.

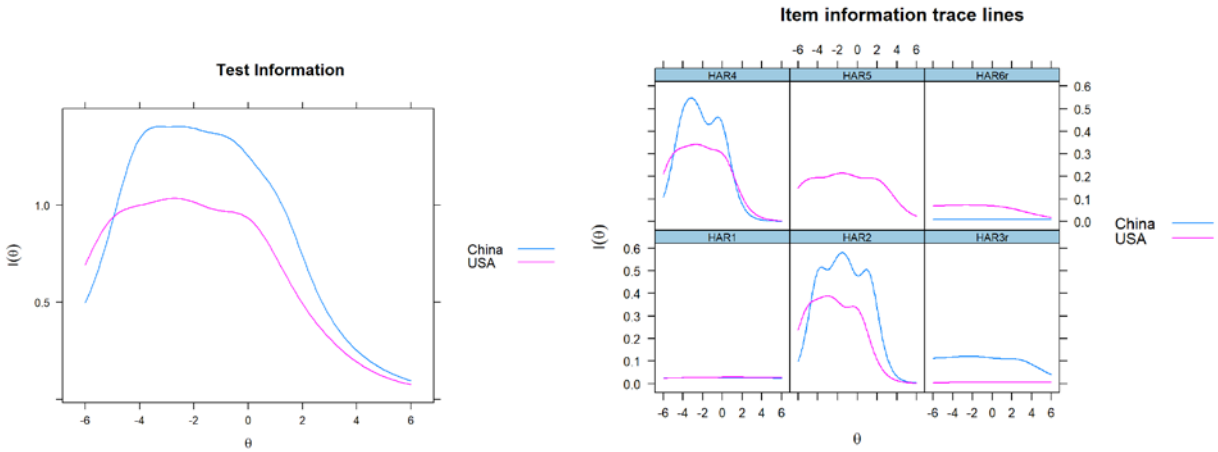

Figure 17. Scale information and item trace plots for the HAR scale.

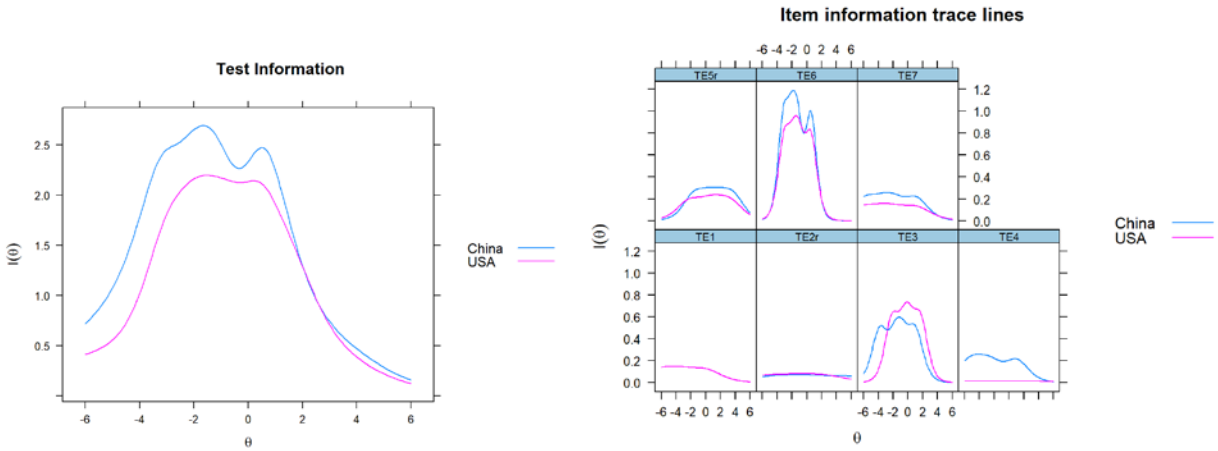

Figure 18. Scale information and item trace plots for the TE scale.

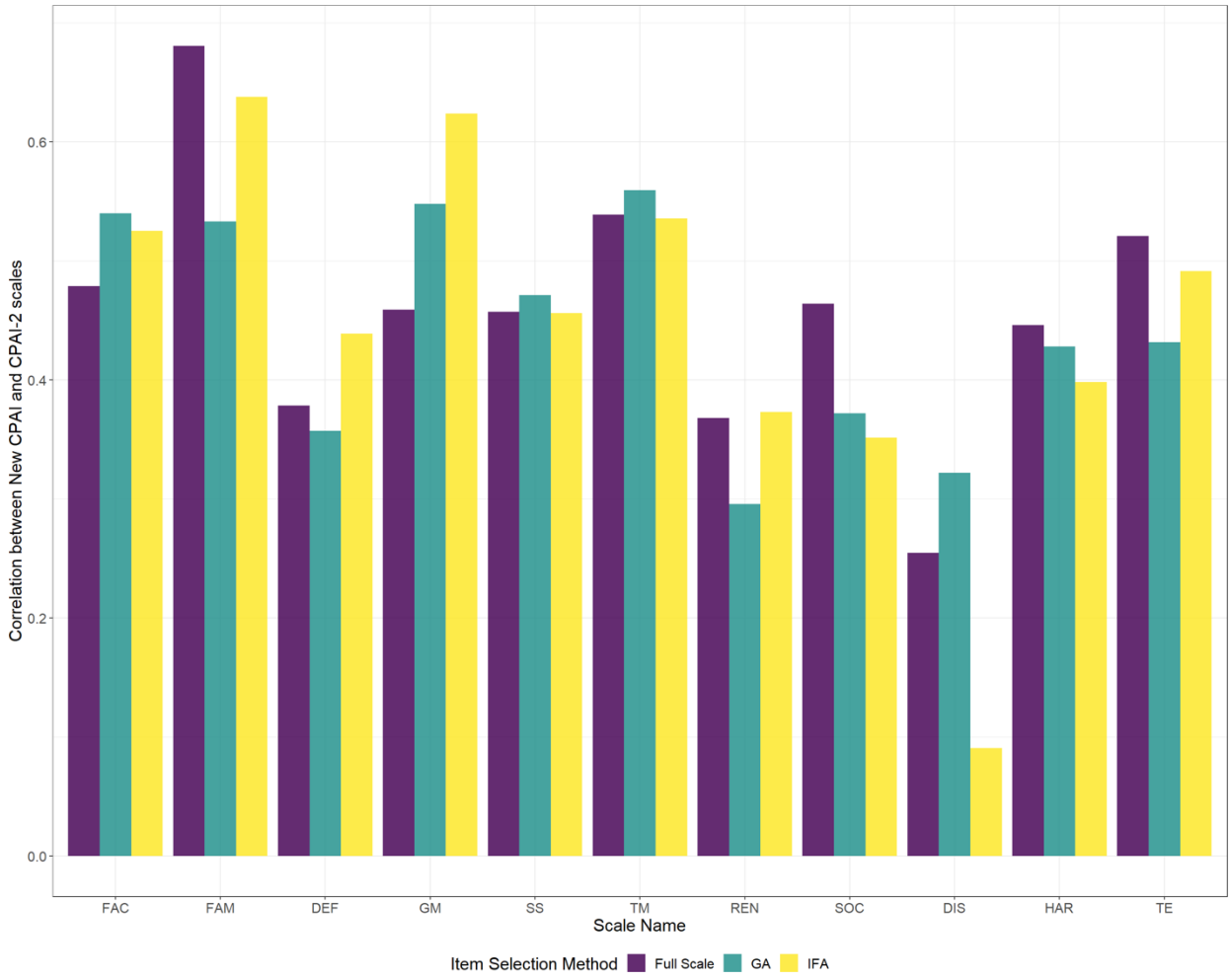

Figure 19. Comparing correlations between New CPAI scales and equivalent CPAI-2 factors using the full scale and the two short scales.

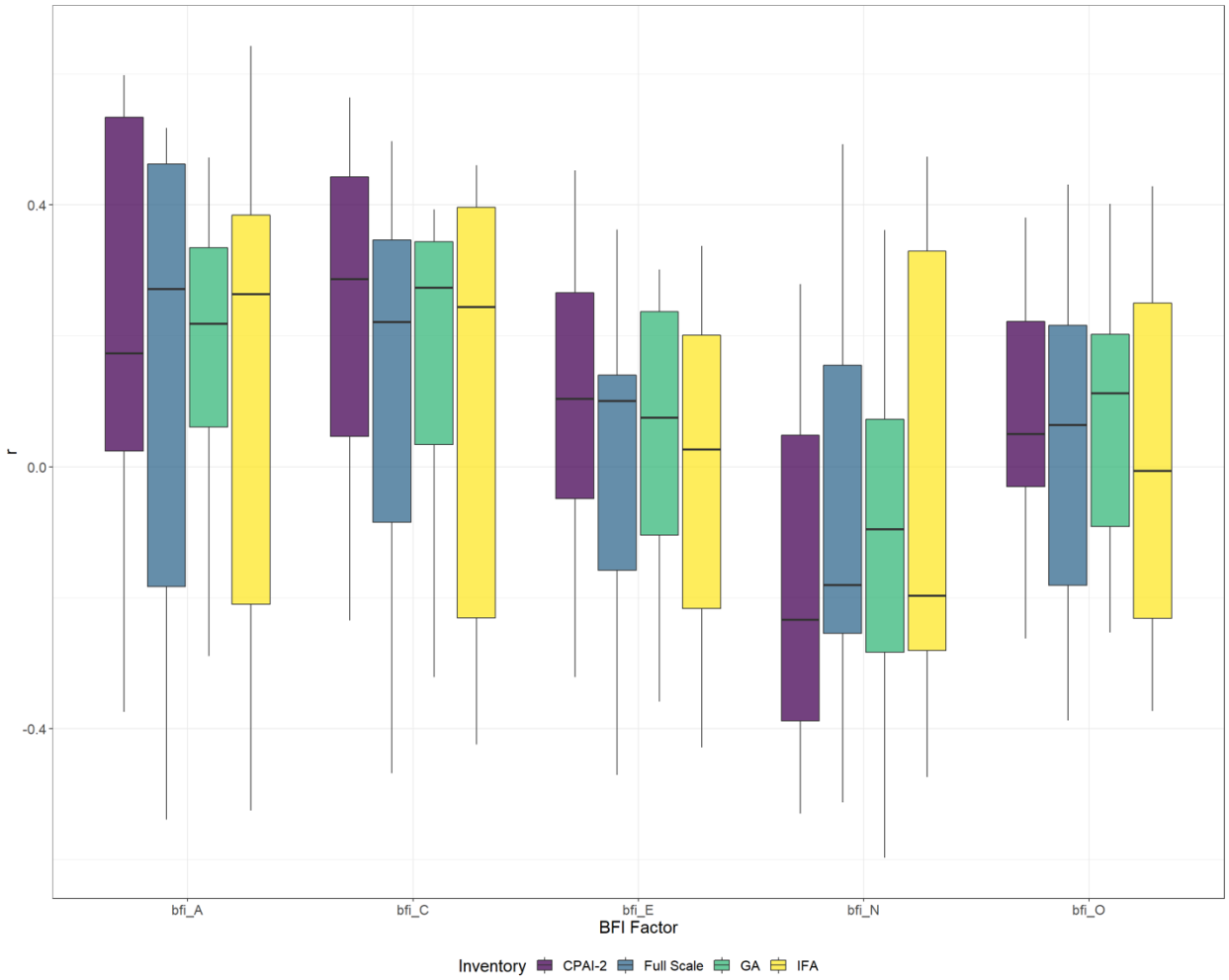

Figure 20. Comparing distributions of correlations between BFI factors and New CPAI and CPAI-2 scales using the full scale and the two short scales.

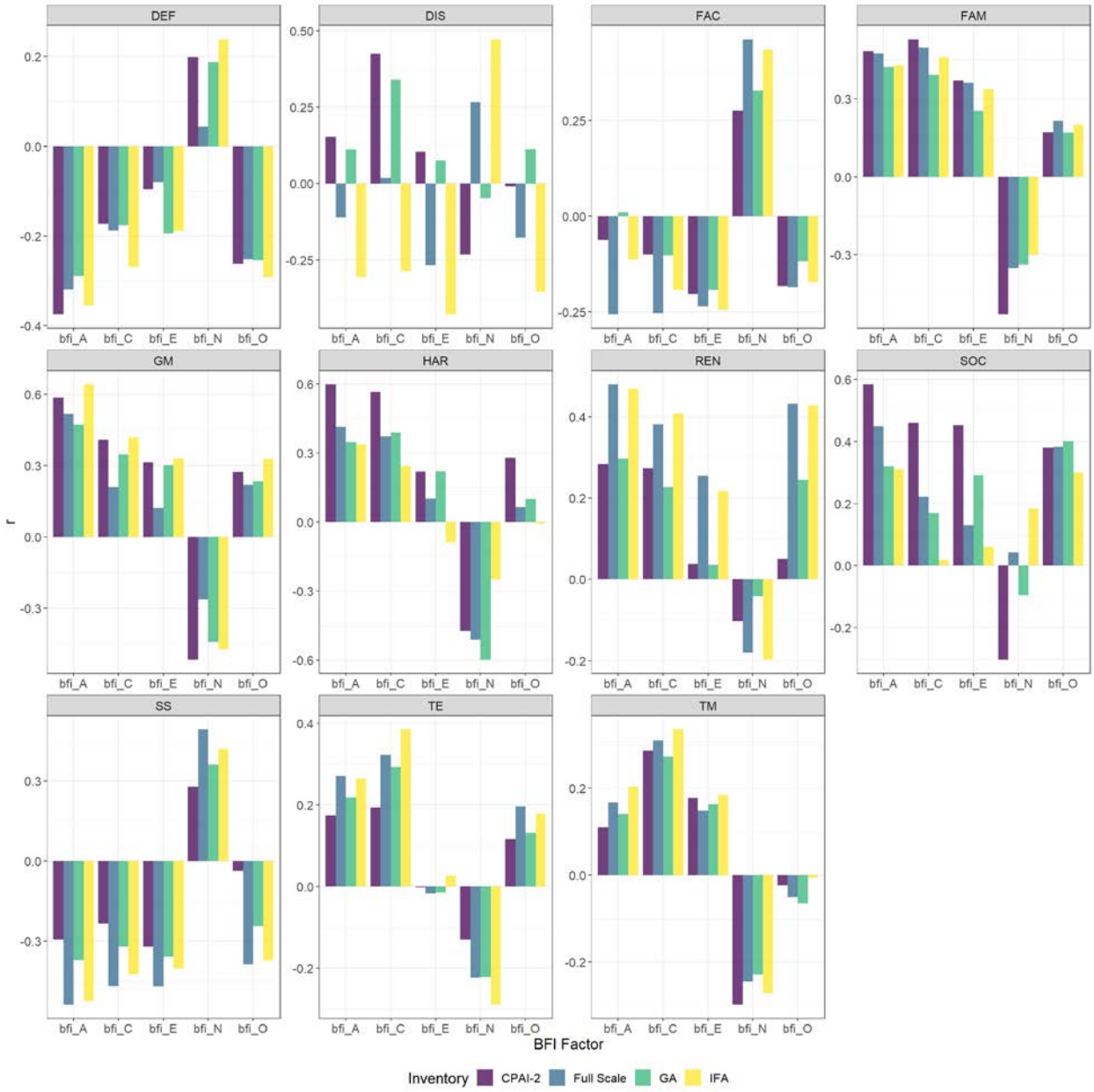

Figure 21. Comparing correlations between BFI factors and New CPAI and CPAI-2 scales using the full scale and the two short scales.

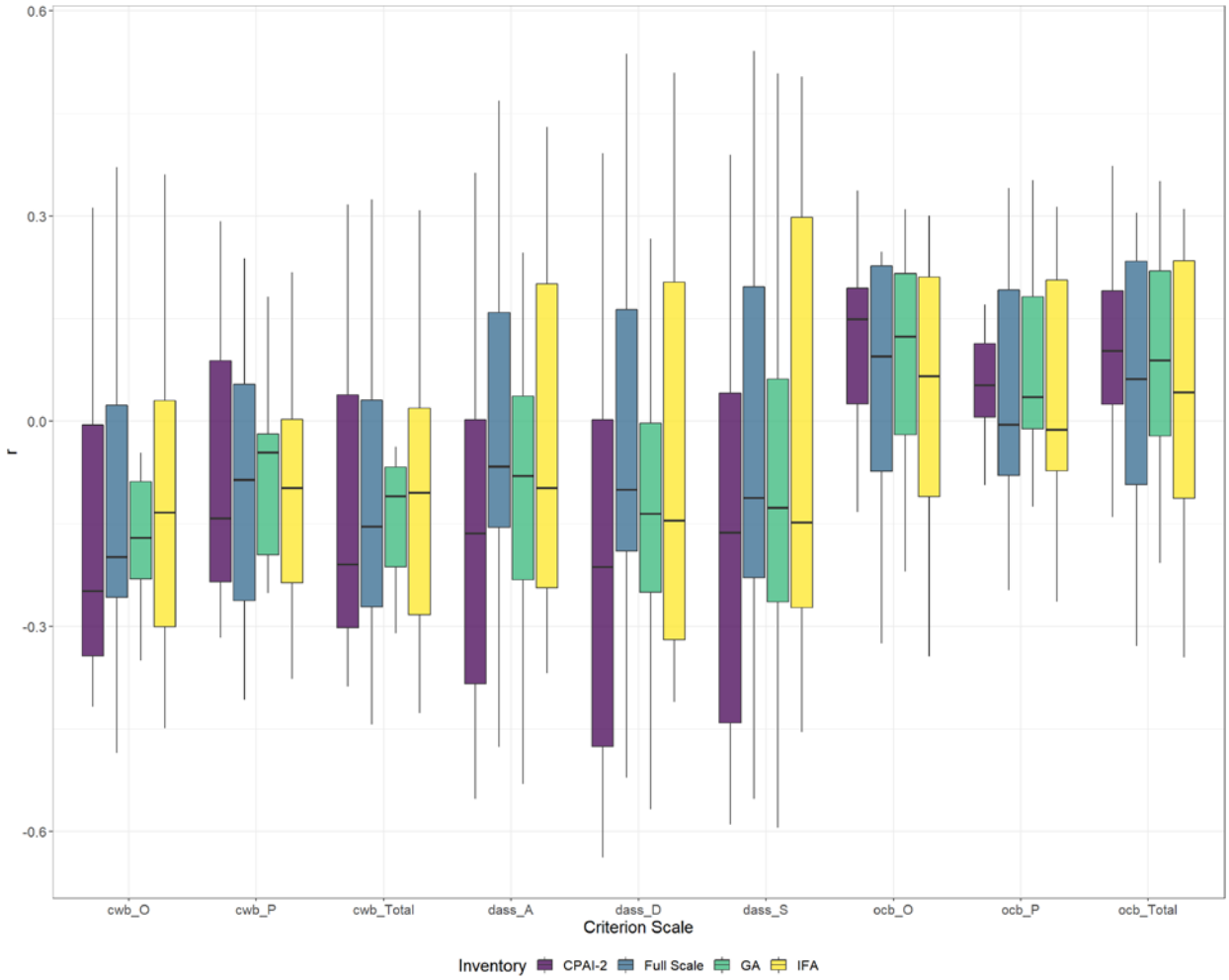

Figure 22. Comparing distributions of correlations between criterion scales and New CPAI and CPAI-2 scales using the full scale and the two short scales.

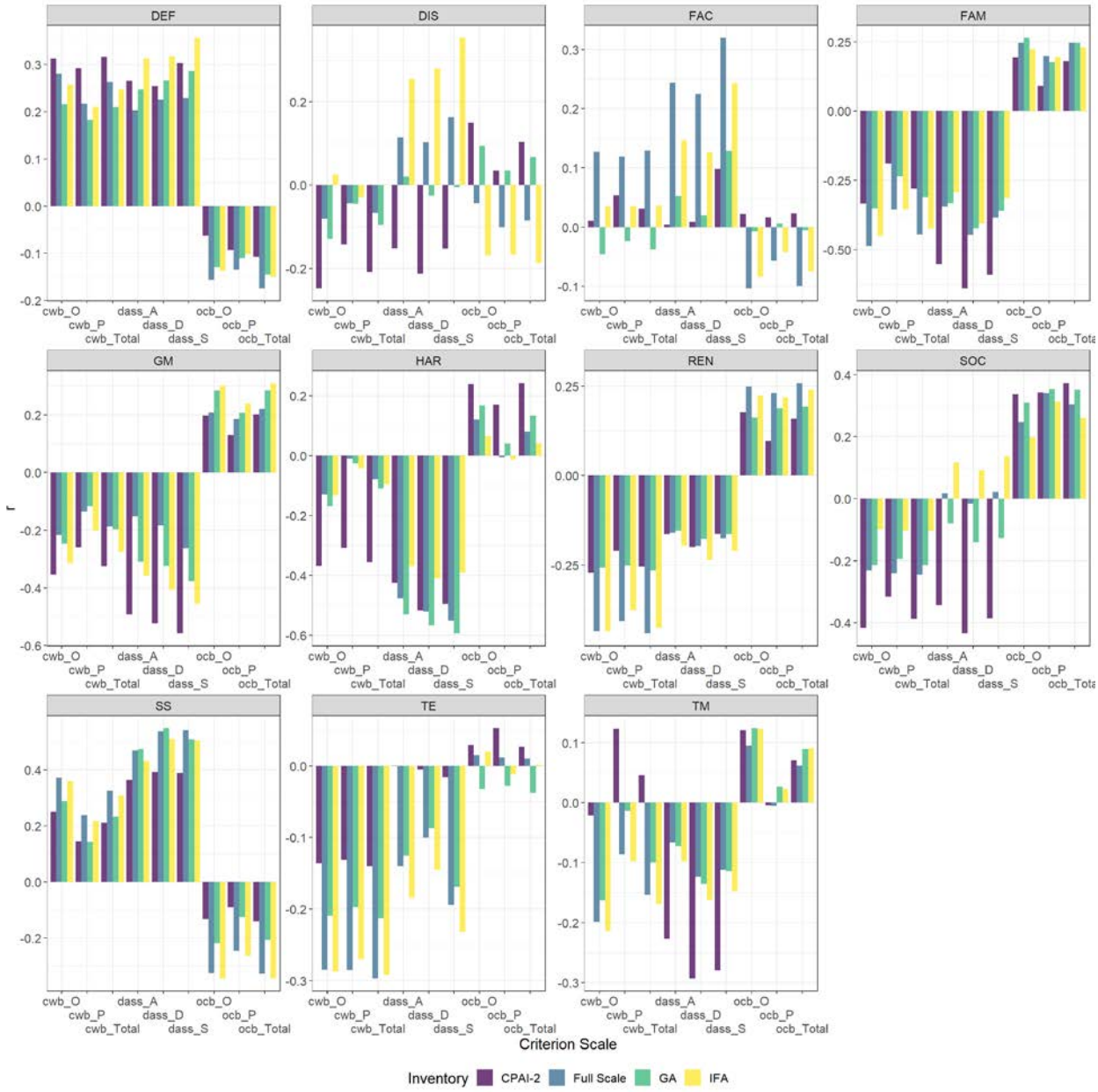

Figure 23. Comparing correlations between criterion scales and all New CPAI and CPAI-2 scales using the full scale and the two short scales.

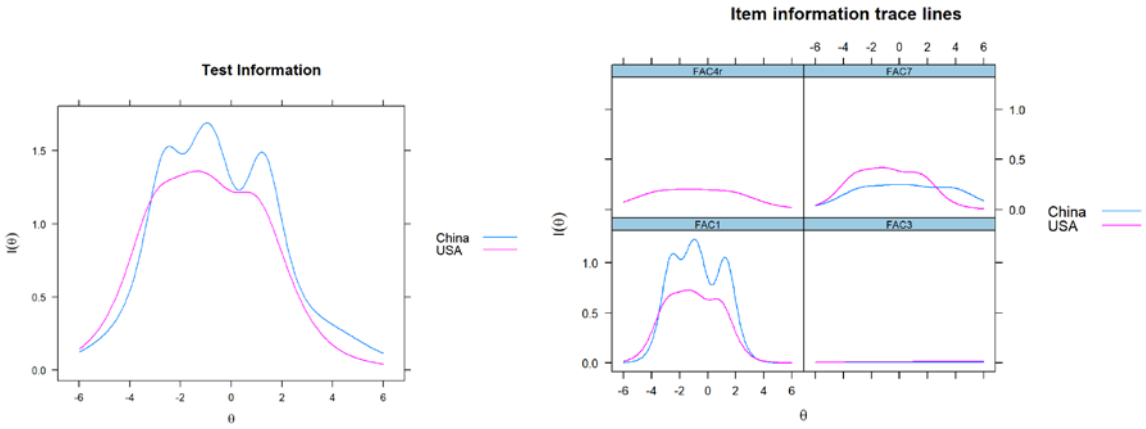

Figure 24. Scale information and item trace plots for FAC items selected using the IFA approach.

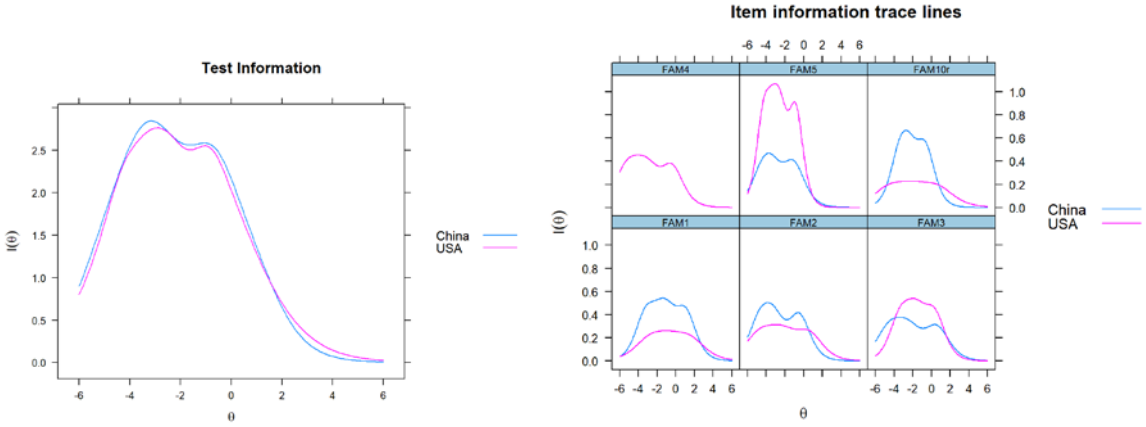

Figure 25. Scale information and item trace plots for FAM items selected using the IFA approach.

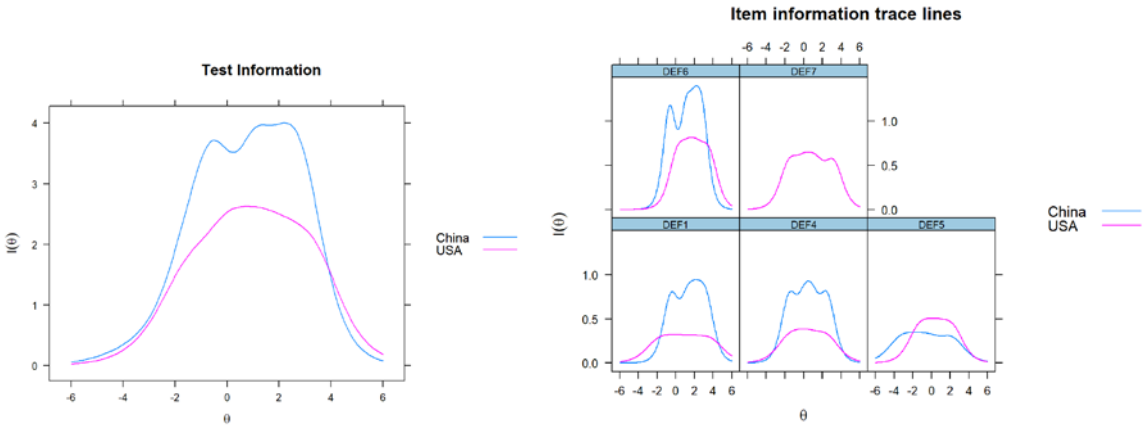

Figure 26. Scale information and item trace plots for DEF items selected using the IFA approach.

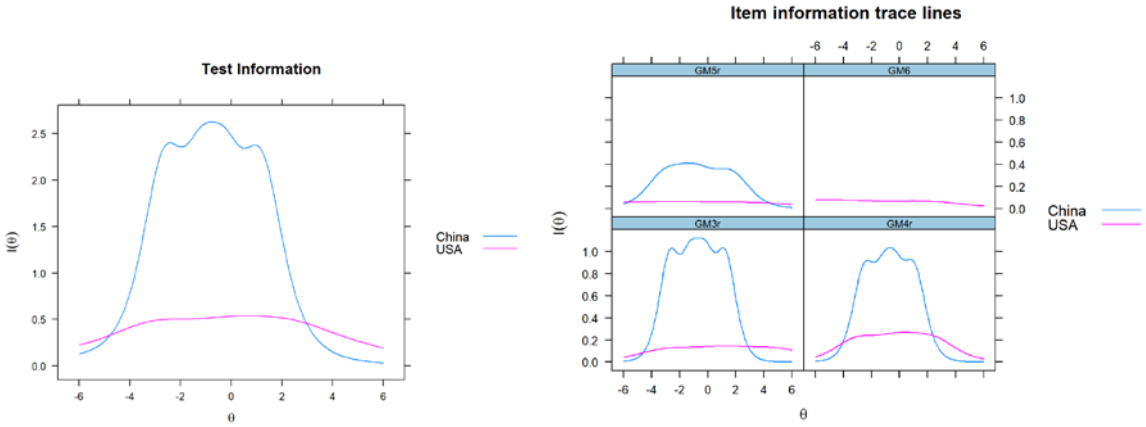

Figure 27. Scale information and item trace plots for GM items selected using the IFA approach.

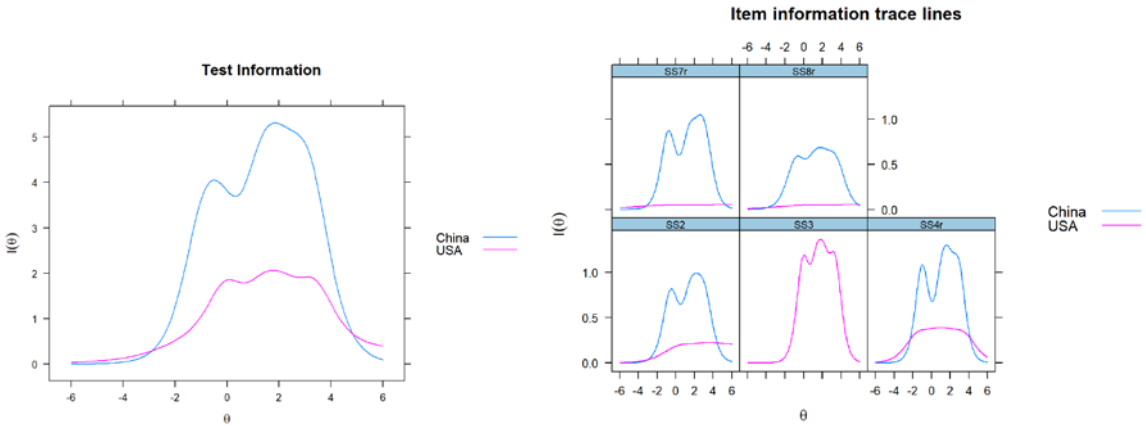

Figure 28. Scale information and item trace plots for SS items selected using the IFA approach.

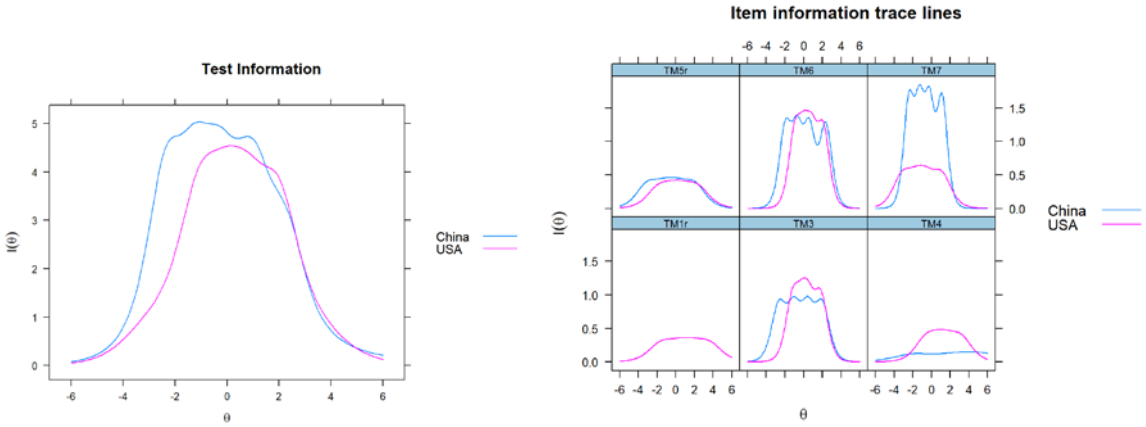

Figure 29. Scale information and item trace plots for TM items selected using the IFA approach.

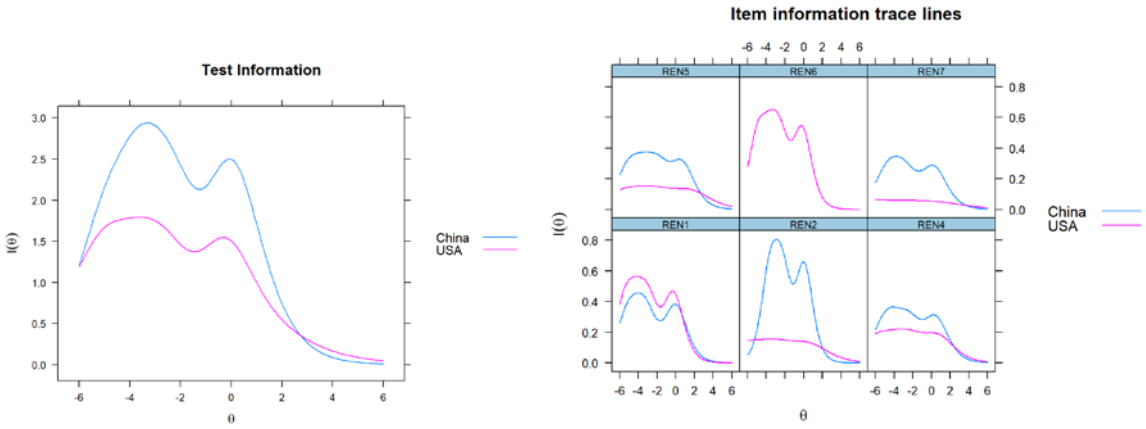

Figure 30. Scale information and item trace plots for REN items selected using the IFA approach.

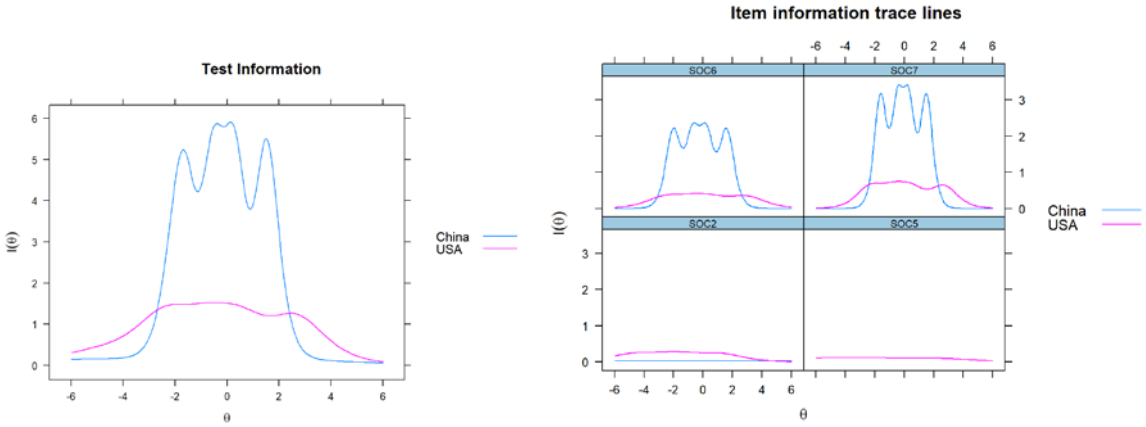

Figure 31. Scale information and item trace plots for SOC items selected using the IFA approach.

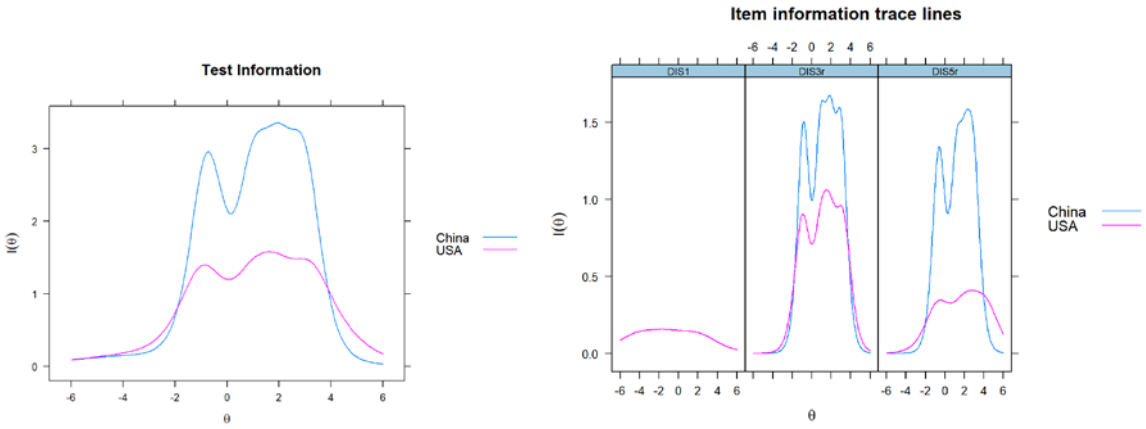

Figure 32. Scale information and item trace plots for DIS items selected using the IFA approach.

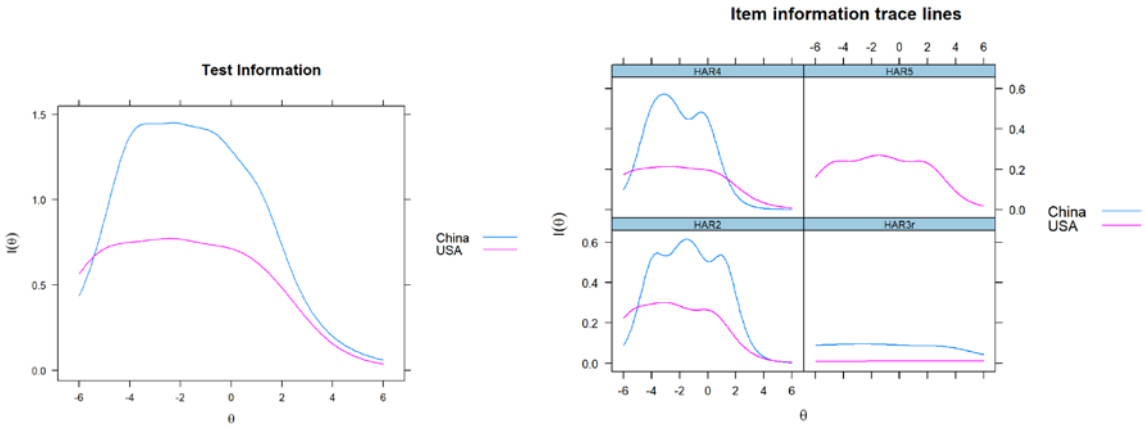

Figure 33. Scale information and item trace plots for HAR items selected using the IFA approach.

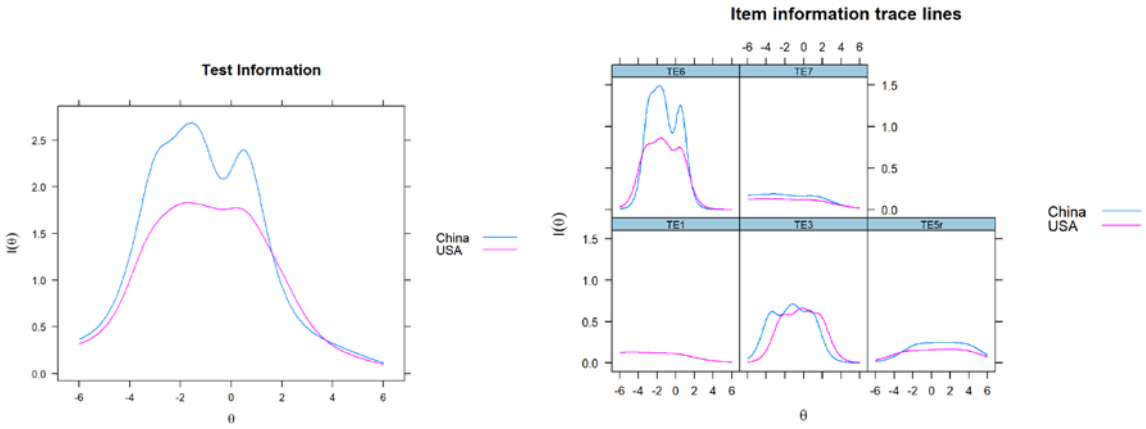

Figure 34. Scale information and item trace plots for TE items selected using the IFA approach.

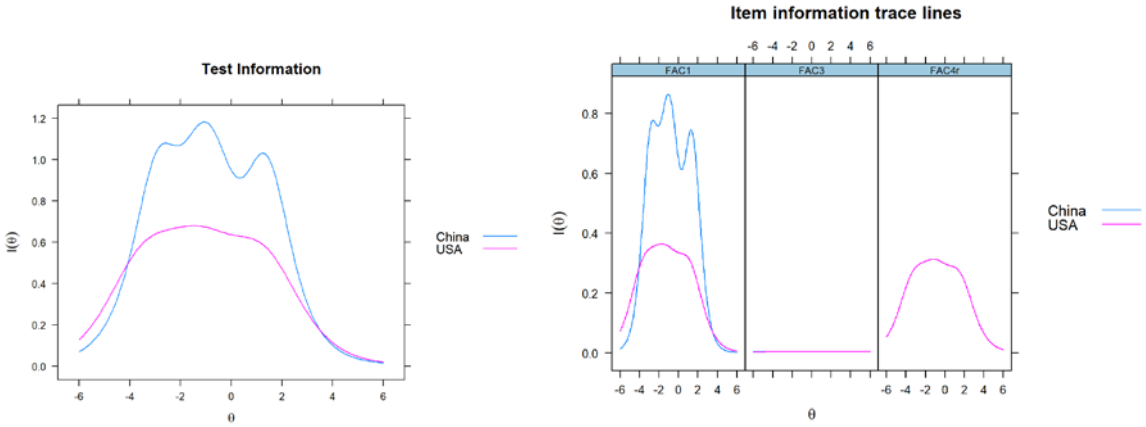

Figure 35. Scale information and item trace plots for FAC items selected using the GA approach.

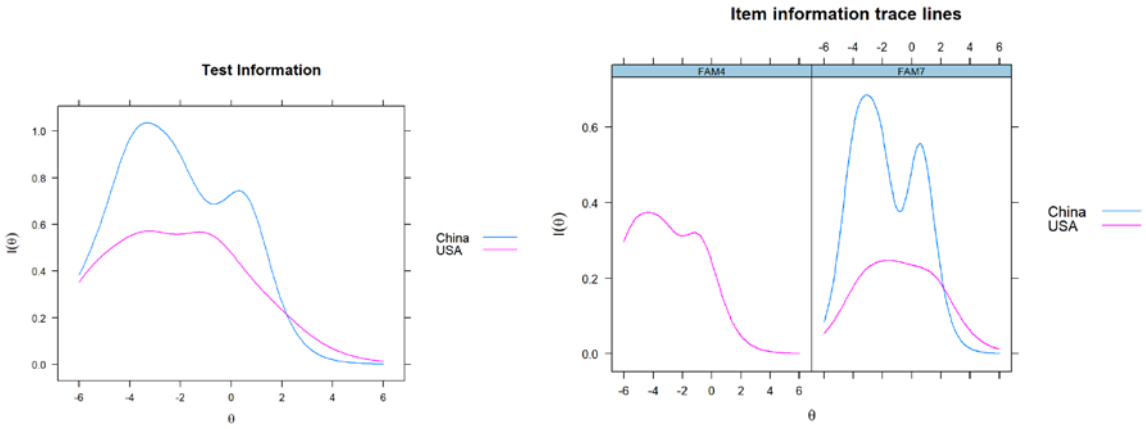

Figure 36. Scale information and item trace plots for FAM items selected using the GA approach.

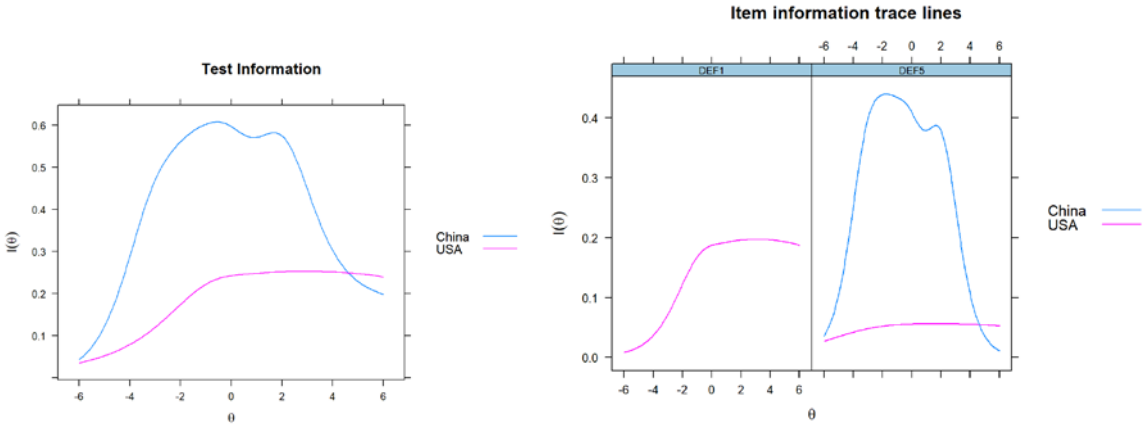

Figure 37. Scale information and item trace plots for DEF items selected using the GA approach.

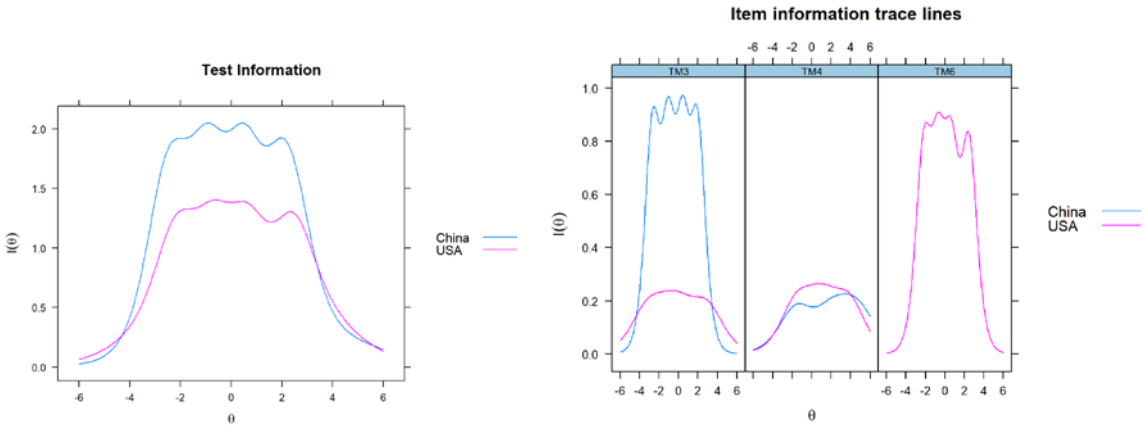

Figure 38. Scale information and item trace plots for TM items selected using the GA approach.

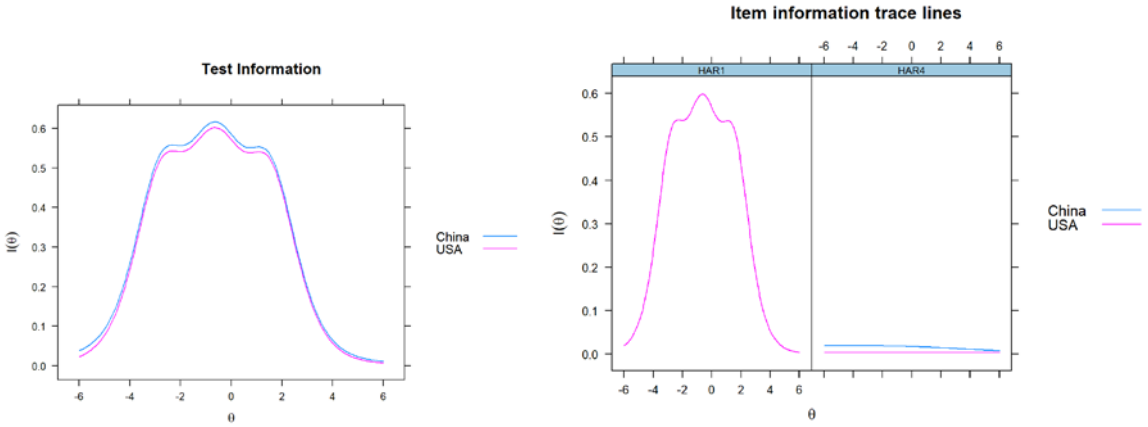

Figure 39. Scale information and item trace plots for HAR items selected using the GA approach.

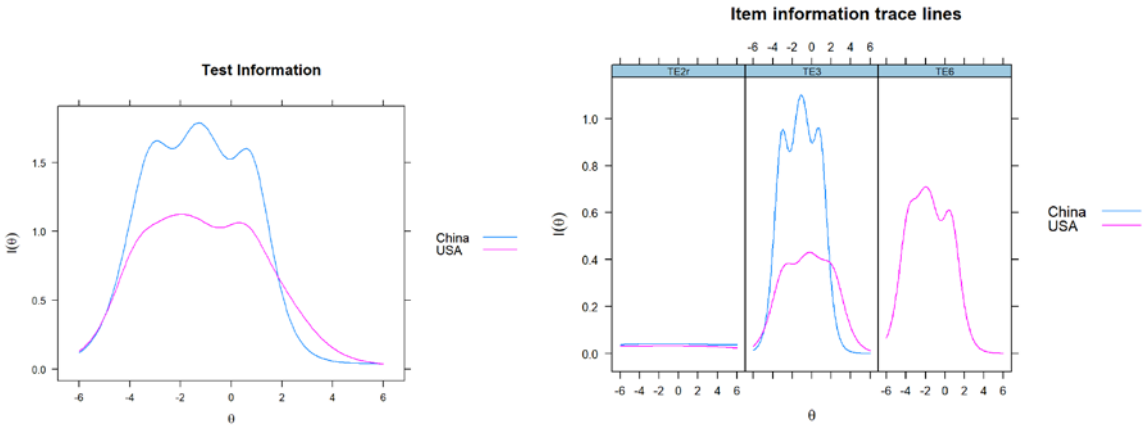

Figure 40. Scale information and item trace plots for TE items selected using the GA approach.

Appendix C:
Item writing worksheet

BUILDING A PERSONALITY TEST

Personality is a construct that describes a person's unique temperament and character based on a combination of **characteristics**. Personality tests are made of **scales** that assess these characteristics. Each scale is constructed using a set of **items** or statements that refer to behaviors or attitudes that reflect those personality characteristics. For example, to find out how extraverted or introverted (the characteristic) a person is, a personality test with an Extraversion-Introversion scale could include the following item: *I am very talkative when I am with a group of people.*

A person taking this personality test would then respond based on how much they agree or disagree with that statement. This response is usually based on a 5-point Likert type scale that ranges from (1) Strongly Disagree to (5) Strongly Agree (see below for an example). The person's responses would then be counted up and their score will then be used to determine their standing on each characteristic. Going back to our previous example, let's say a high score on the Extraversion-Introversion scale means that the person is more extraverted. If a person's total score is low on that scale, that is, they responded "strongly disagree" to most items, we can safely say that they tend to be more introverted.

Let's take a look at our example again. The Extraversion-Introversion scale is described by one's social orientation and styles of interaction. A high score on this scale indicates someone who is very sociable and socially comfortable. A low score indicates someone who prefers to be quiet and solitary.

Our sample item for assessing this scale is:

I am very talkative when I am with a group of people.

These are the possible responses a person can choose from when responding to that item:

Your task is to come up with some items for each scale in this personality test. There are 13 scales that each measure a unique characteristic. Shoot for about *5 to 10 items* for each of the 13 scales.

On the following pages, you will be presented with the name of the scale, a description of characteristic the scale is trying to assess, and what a high score or a low score on that scale means. Write the items in such a way that a person taking the test would be able to respond to it using the 5-point Likert type scale described above and pictured below for your reference.

I am very talkative when I am with a group of people.

Thank you for agreeing to help build this list of items! Your input will be very valuable for the study of personality and psychological science.

Scale Name	Scale Description	Scale dimensions
Face	One's concern for maintaining a proper reputation and image in social interactions. Focuses excessive attention on social recognition, concern for maintaining self-respect in social relationships. Face is a dominant concept in interpreting and regulating social behavior in Chinese culture.	*High:* Places a lot of concern for one's reputation and image. *Low:* Does not care much about reputation or social-image.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Family Orientation	One's sense of family solidarity and responsibility. These family ties provide emotional and financial security and support, especially in Asian or collectivistic cultures.	***High:*** Values family bonding and strong family ties. ***Low:*** Cares very little about responsibility to family and family solidarity.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Defensiveness	One's use of defense mechanisms, such as self-protective rationalization, externalization of blame, self-enhancement, and belittling of others' achievements. A mild degree is accepted as a healthy protective mechanism against defeatism and depression.	***High:*** Uses protective defense mechanisms to feel better about oneself. ***Low:*** Does not engage in a great deal defensive behaviors or thoughts.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Graciousness versus Meanness	Extent to which one is kind and broad-minded in their dealings with others.	***High:*** Bears no grudges, treats others leniently. ***Low:*** Overly critical of others, retaliatory, and calculating.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Self vs. Social Orientation	Degree of enthusiasm for teamwork and willingness to contribute to the collective over the individual goals.	***High:*** Independent and unwilling to join cooperative activities. ***Low:*** Collectivistic orientation, a team player.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Traditionalism versus Modernity	The degree of individual modernization as an indication of one's responses to societal modernization. This scale covers attitudes toward traditional cultural beliefs and values in the areas of family relationships, filial piety, social rituals, and chastity that are emphasized in Chinese societies.	***High:*** Endorses traditional beliefs, customs, and values. ***Low:*** Challenges traditional ideas, endorses individual freedom.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Renqing (Relationship Orientation)	One's adherence to the cultural norms of reciprocal interactions such as courtesy rituals, exchanging resources, maintaining and utilizing useful ties, and nepotism.	***High:*** Endorses cultural norms of courtesy and maintains useful ties. ***Low:*** Challenges cultural relationship norms and ideologies.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Social Sensitivity	The extent to which an individual is empathic and sensitive to how others feel.	***High:*** Sensitive to how others feel and react. ***Low:*** Has little interest in other people's feelings.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Discipline	How rigid and disciplined one is as opposed to being adaptable, flexible, and carefree.	***High:*** Rigid and disciplined. ***Low:*** Flexible and adaptable.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Harmony	The degree of one's inner peace of mind, contentment, and interpersonal relations with others. The avoidance of conflict and maintenance of equilibrium are considered important virtues in Asian traditions and collectivistic cultures.	***High:*** Maintains inner peace and contentment and avoids conflict and competition. ***Low:*** Often gets into conflict with others and faces a great deal of inner turmoil.

Write your items in the space below:

Scale Name	Scale Description	Scale dimensions
Thrift versus Extravagance	One's tendency to save rather than waste and one's carefulness in spending. Thrift is one of the basic traditional Confucian Chinese values, and the characteristic of thrift versus extravagance is an indicator of the social response to rapid economic development and increasing materialism.	***High:*** Endorses traditional value of frugality. ***Low:*** Endorses materialistic tendencies and high consumption.

Write your items in the space below: