The personality of U.S. states: Stability from 1999 to 2015 ☆

Lorien G. Elleman [a,*], David M. Condon [b], Sarah E. Russin [a], William Revelle [a]

[a] Department of Psychology, Northwestern University, Evanston, IL 60208, United States
[b] Department of Medical Social Sciences, Northwestern University, Chicago, IL, United States

ARTICLE INFO

ABSTRACT

Researchers have shown an interest in the aggregated Big Five personality of U.S. states, but typically they have relied on scores from a single sample (Rentfrow, Gosling, & Potter, 2008). We examine the replicability of U.S. state personality scores from two studies (Rentfrow et al., 2008; Rentfrow, Gosling, Jokela, & Stillwell, 2013) across a total of seven samples, two of them new. Same-trait correlations across samples are, on average, positive for all five traits, indicating score agreement. Additionally, three traits (Conscientiousness, Neuroticism, and Openness) show strongly consistent patterns of correlations with sociodemographic variables across samples. We find rank order stability in state personality scores for a 16-year period (1999–2015).

1. Introduction

The beginning of the twenty-first century has seen an explosion of interest concerning geographical variation in personality within the United States. Before the modern era of the internet, there were a few studies that examined aggregate psychological differences by U.S. cities or regions (e.g., Krug & Kulhavy, 1973; Thorndike, 1939). However, with the widespread adoption of the internet in the U.S., several psychology labs have collected samples of hundreds of thousands of participants across the country via online personality assessments (e.g., Revelle, Wilt, & Rosenthal, 2010; Revelle et al., 2016; Srivastava, John, Gosling, & Potter, 2003). These samples, although not representative of the U.S. population, are more diverse than traditional methods of data collection (Gosling, Vazire, Srivastava, & John, 2004). They also have enough statistical power for analyses to detect small effects between a large number of regional groups. These online assessments typically use self-report personality assessment models based on the Big Five (Goldberg, 1990), a widely-accepted taxonomy that organizes most individual differences into five broad traits: Conscientiousness, Agreeableness, Neuroticism (sometimes referred to by its polar opposite, Emotional Stability), Openness (sometimes called Intellect), and Extraversion.

Rentfrow, Gosling, and Potter (2008) were the first to showcase one of these large samples ($n = 619,397$) in a landmark study of U. S. regional differences of Big Five personality. Their study was the first to aggregate individual Big Five personality scores into mean state scores for all 50 states and Washington, DC. They were also first to publish the state scores for each trait, in the form of ranks and standardized scores. These published scores have proven to be of long-lasting utility to researchers interested in U.S. regional personality differences. Collecting such a massive sample is several orders of magnitude easier than in the last century, but it is by no means a trivial endeavor. Thus, many researchers have leveraged these published data to correlate U.S. state scores with state-level sociodemographics, such as chronic disease (Pesta, Bertsch, McDaniel, Mahoney, & Poznanski, 2012), obesity (McCann, 2011), income inequality (de Vries, Gosling, & Potter, 2011), and the severity of state governmental punishment (Harrington & Gelfand, 2014). Rentfrow et al. (2008) correlated state scores with sociodemographics and found some results that were congruent with their hypotheses (e.g., Openness was positively related to liberal values) and some that were unexpected (e.g., Extraversion was positively related to murder per capita).

Studies that have used the state scores data from Rentfrow et al. (2008) have assumed that these state scores were representative of

the actual personality scores of the states' residents. For example, these studies assumed that the Extraversion score for Oklahoma accurately represented the mean Extraversion score of all Oklahomans. This assumption could be problematic for at least three reasons. First, these state scores, although based on many participants, are not immune to sampling bias. Idiosyncratic methods of participant selection could lead to a lack of replication in other samples. Second, even if the original scores were accurate, the personality of some states' residents may have changed since the original study's data were collected (1999–2005). Third, Rentfrow et al. (2008) measured personality with the Big Five Inventory (BFI; John & Srivastava, 1999). The ranks and standardized scores from Rentfrow et al. (2008) may not generalize across other measures of the Big Five. Therefore, it is useful to examine the extent to which the state-level scores of Rentfrow et al. (2008) will replicate, and to estimate the effect of disagreement attributable to differences in participant recruitment methods, change over time, and Big Five measures. Additionally, because these state scores are often reused in other studies, it is critical to determine the extent to which correlations between state personality and sociodemographics replicate in spite of differences in samples.

A study dedicated to replicating Rentfrow et al. (2008) has not yet been reported. Rentfrow, Gosling, Jokela, and Stillwell (2013) compared state scores across five samples, including the sample from Rentfrow et al. (2008). Details from this effort were brief because the focus of the paper concerned the personality profiles of broad regions of the U.S. Findings were not thoroughly discussed, and many of the results were relegated to the supplemental materials. However, their analyses indicated that "there were no clear or consistent statewide differences in any of the scale properties," and state scores were "reliable and generalizable" (Rentfrow et al., 2013, p. 1003). The study also found that in general, the same traits correlated with the same sociodemographic variables at similar magnitudes across the five samples.

The current study used the five samples reported from Rentfrow et al. (2013) and added two new large samples. The goals of the study were as follows: One, for a point of comparison with Rentfrow et al. (2013), determine the extent to which the two replication samples were representative of U.S. states (Section 3.1). Two, across all samples, estimate the reliability of state score differences by evaluating their intraclass correlations (Section 3.2). Three, estimate the effect size of same-trait convergent correlations across all samples for each Big Five trait (Section 3.3). Within this goal, estimate the separate effects of three possible sources of attenuation: differences due to recruitment methods (Section 3.3.1), time of data collection (Section 3.3.2), and personality inventories (Section 3.3.3). Four, from these estimates, determine whether the personality of U.S. states had maintained rank order stability (i.e., relative to each other, states' personalities did not change) from 1999 to 2015 (Section 3.3.4). And finally, determine the replicability of correlations between state-level personality scores and sociodemographic variables (Section 3.4).

2. Methods

2.1. Samples 1–5

Samples 1–5 were originally analyzed in Rentfrow et al. (2013). The samples were collected during different time periods, as part of different research projects, using different personality inventories (Table 1). All five samples had aggregate measures of personality based on Likert-type scales for the 48 contiguous states, as well as Washington, DC. In total, the samples were collected over an 11-year period (1999–2010), with a range in sample size from

Table 1
Summary description of samples.

Sample	Participant count	Personality inventory	Research project	Time period
Sample 1	612,140	44-item BFI	Gosling-Potter	1999–2005
Sample 2	507,987	44-item BFI	Gosling-Potter	2005–2009
Sample 3	145,307	10-item TIPI	Rentfrow-Potter	2002–2009
Sample 4	312,568	20-item NEO	MyPersonality	2008–2010
Sample 5	18,182	10-item TIPI	CCAP	2007–2008
SAPA2010	81,538	100-item BFFM	SAPA	2006–2010
SAPA2015	134,858	100-item BFFM	SAPA	2010–2015

18,182 to 612,140. Samples 1–4 were online personality assessments that used self-selecting participant recruitment. Sample 5 was an online assessment that used a recruitment method similar to random digit-dialing to select a representative sample of registered voters. A more detailed summary of Samples 1–5 can be found in Rentfrow et al. (2013), where each sample is referred to by the same name used in this study.

Through correspondence with Rentfrow, we received participant counts and unadjusted mean scores for states, for each sample. Other data on the five samples, such as interclass correlations, were collected from Rentfrow et al. (2013) and the supplemental materials.

2.2. SAPA samples

The last two samples were from the Synthetic Aperture Personality Assessment (SAPA) project, an online non-commercial personality assessment (https://sapa-project.org; Revelle et al., 2016). For their participation, participants received feedback concerning their personality. Each sample covered an approximate five-year period of time and was named for the last year in which data were collected. The SAPA2010 sample was collected from April 2006 to August 2010. The SAPA2015 sample was collected from August 2010 to December 2015 (Table 1).

2.2.1. Participants

Participants were screened to ensure that entries beyond their first were not included in the analysis. Duplicate entries taken in a single internet browser session were removed. Participants who reported having previously taken the assessment also were excluded. Since this study was concerned with state-level analysis, it was also necessary to remove participants who reported not being from one of the 50 U.S. states or Washington, DC.

After these screening procedures, there were 81,538 participants in the SAPA2010 sample (70% female), and 134,858 participants in the SAPA2015 sample (66% female). The median age for the SAPA2010 and SAPA2015 samples was 23 years (Median Absolute Deviation [MAD] = 7.4) and 22 years (MAD = 5.9), respectively. Of the 81,532 SAPA2010 participants who reported their race or ethnicity, 77% were white, 8% were African American, 6% were Hispanic, 4% were Asian, 1% were Native Alaskan/American/Hawaiian, and 4% reported being "Other." Of the 132,838 SAPA2015 participants who reported their race or ethnicity, 67% were white, 10% were African American, 9% were Hispanic, 5% were Asian, 1% were Native Alaskan/American/Hawaiian, 6% were more than one race or ethnicity, and 1% reported being "Other."

Concerning educational attainment, 40% of the SAPA2010 sample reported being an undergraduate at the time of assessment, while 28% had attained at least a bachelor's degree. In the SAPA2015 sample, 51% were current undergraduates, while 27% had attained at least a bachelor's degree.