

# **Defining collectivism as duties in relationships rather than generalized warmth improves predictive validity in 100 cultures**

---

In the format provided by the  
authors and unedited

## Supplementary Information

### Table of Contents

|    |                                                                                                                                                    |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------|
|    | <a href="#"><u>Section 1</u></a> : Does Responsibilism Predict Actual Behavior?                                                                    |
|    | <a href="#"><u>Section 2</u></a> : Convergent Validity and Societal Markers                                                                        |
| 5  | <a href="#"><u>Section 3</u></a> : Demographic Questions and Cultural Differences Controlling for Demographics                                     |
|    | <a href="#"><u>Section 4</u></a> : Sample Selection                                                                                                |
|    | <a href="#"><u>Section 5</u></a> : Scale Development and Translation                                                                               |
|    | <a href="#"><u>Section 6</u></a> : Scale Refinement and Item Selection                                                                             |
|    | <a href="#"><u>Section 7</u></a> : Component Scores and Factor Structure                                                                           |
| 10 | <a href="#"><u>Section 8</u></a> : Responsibilism Varies Significantly Between Cultures                                                            |
|    | <a href="#"><u>Section 9</u></a> : Responsibilism Can Measure Individual Differences and Cultural Differences (Cross-Level Measurement Invariance) |
|    | <a href="#"><u>Section 10</u></a> : Scale Equivalence Across Cultures: Measurement Invariance                                                      |
|    | <a href="#"><u>Section 11</u></a> : Scale Reliability                                                                                              |
| 15 | <a href="#"><u>Section 12</u></a> : Survey of Researchers                                                                                          |
|    | <a href="#"><u>Section 13</u></a> : Testing Earlier Collectivism Scales for Divergent Validity                                                     |
|    | <a href="#"><u>Section 14</u></a> : Is Responsibilism Related to Relational Mobility and Tightness?                                                |
|    | <a href="#"><u>Section 15</u></a> : Data-Quality Check: Latitude                                                                                   |
|    | <a href="#"><u>Section 16</u></a> : Data-Quality Check: Careless Responding                                                                        |
| 20 | <a href="#"><u>Section 17</u></a> : Data-Quality Check: Survey Duration                                                                            |
|    | <a href="#"><u>Section 18</u></a> : Accounting for Response Styles                                                                                 |
|    | <a href="#"><u>Section 19</u></a> : Estimating Culture Means                                                                                       |
|    | <a href="#"><u>Section 20</u></a> : Prior Collectivism Scales Mix Warm Fuzzy Intuitions and Responsibilistic Intuitions                            |
|    | <a href="#"><u>Section 21</u></a> : Rating Self vs. Others                                                                                         |
| 25 | <a href="#"><u>Section 22</u></a> : Did the Pandemic Affect the Results?                                                                           |
|    | <a href="#"><u>Section 23</u></a> : Longitudinal Test-Retest Reliability                                                                           |
|    | <a href="#"><u>Section 24</u></a> : Definition of West, Sometimes West, and Non-West                                                               |
|    | <a href="#"><u>Section 25</u></a> : Cultures Versus Countries                                                                                      |
|    | <a href="#"><u>Section 26</u></a> : Scale Creation: Avoiding English-Centric Bias                                                                  |
| 30 | <a href="#"><u>Section 27</u></a> : What About Hofstede’s Cultural Dimensions?                                                                     |
|    | <a href="#"><u>Section 28</u></a> : Social Class Analysis                                                                                          |
|    | <a href="#"><u>Section 29</u></a> : Percentage of the World Population Covered                                                                     |
|    | <a href="#"><u>Section 30</u></a> : Accounting for Response Variability Did Not Improve Predictive Power                                           |
|    | <a href="#"><u>Section 31</u></a> : Spatial Auto-Correlation: Controlling for Latitude and Longitude                                               |
| 35 | <a href="#"><u>Section 32</u></a> : Scale Refinement Using a Holdout Sample                                                                        |
|    | <a href="#"><u>Section 33</u></a> : Research Articles Mentioning “Collectivism” Increased 190% in the 1990s                                        |
|    | <a href="#"><u>Section 34</u></a> : Robustness to Even Sample Sizes and Samples of 40 Participants per Culture                                     |
|    | <a href="#"><u>Section 35</u></a> : What Counts as a Close Relationship?                                                                           |
|    | <a href="#"><u>Section 36</u></a> : Evidence That People in Individualistic Cultures Tend to Agree More with Warm                                  |
| 40 | Fuzzy Collectivism Items                                                                                                                           |
|    | <a href="#"><u>Section 37</u></a> : The Pipeline Project Two Replication                                                                           |
|    | <a href="#"><u>Section 38</u></a> : Checking Assumptions of Statistical Tests                                                                      |
|    | <a href="#"><u>Fig. S1</u></a> . Correlations between responsibilism and external markers of collectivism.                                         |
|    | <a href="#"><u>Fig. S2</u></a> . Correlations between responsibilism and external markers of collectivism.                                         |
| 45 | <a href="#"><u>Fig. S3</u></a> . Responsibilistic cultures have lower relational mobility.                                                         |
|    | <a href="#"><u>Fig. S4</u></a> . Responsibilistic cultures have tighter social norms.                                                              |
|    | <a href="#"><u>Fig. S5</u></a> . Responsibilism ratings for self-report vs. “culture report” in the US and Hong Kong.                              |



|    |                                                                                                                                                                                                 |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    | <a href="#"><u>Fig. S6.</u></a> Responsibilism component scores for family responsibilities.                                                                                                    |
|    | <a href="#"><u>Fig. S7.</u></a> Responsibilism component scores for guarding personal details.                                                                                                  |
|    | <a href="#"><u>Fig. S8.</u></a> Responsibilism component scores for tight friendship.                                                                                                           |
|    | <a href="#"><u>Fig. S9.</u></a> Responsibilism component scores for stranger danger.                                                                                                            |
| 5  | <a href="#"><u>Fig. S10.</u></a> Responsibilism component scores for fitting in.                                                                                                                |
|    | <a href="#"><u>Fig. S11.</u></a> Cultural psychologists' guesses about cultural differences versus actual cultural differences.                                                                 |
|    | <a href="#"><u>Fig. S12.</u></a> Responsibilism went down in China during the COVID-19 pandemic (left), but the change was far smaller than East-West differences (right).                      |
| 10 | <a href="#"><u>Fig. S13.</u></a> Responsibilism went up in India from 2014 before the COVID-19 pandemic (left) to 2021 (middle), but both samples showed higher responsibilism than US (right). |
|    | <a href="#"><u>Fig. S14.</u></a> Responsibilism scores with no sample restrictions (X axis) versus excluding participants with non-local cultural backgrounds (Y axis).                         |
|    | <a href="#"><u>Fig. S15.</u></a> Responsibilism scores without data-quality exclusions (X axis) versus with data-quality exclusions (Y axis).                                                   |
| 15 | <a href="#"><u>Fig. S16.</u></a> Multigroup factor analysis.                                                                                                                                    |
|    | <a href="#"><u>Fig. S17.</u></a> Percentage of psychology articles mentioning "collectivism" from 1950 to 2019.                                                                                 |
|    | <a href="#"><u>Fig. S18A.</u></a> Distribution of correlations of responsibilism with the Global Collectivism Index in 50 random samples of 40 participants per culture.                        |
| 20 | <a href="#"><u>Fig. S18B.</u></a> Distribution of external validity correlations from 50 random samples of 40 participants per culture.                                                         |
|    | <a href="#"><u>Fig. S19.</u></a> Estimates of culture scores for responsibilism across 50 random samples ranging from 5 to 100 participants per culture.                                        |
| 25 | <a href="#"><u>Table S1.</u></a> Responsibilism predicts external markers of collectivism more strongly than previous collectivism scales.                                                      |
|    | <a href="#"><u>Table S2.</u></a> Correlations of collectivism scales and external markers of collectivism controlling for economic development.                                                 |
|    | <a href="#"><u>Table S3.</u></a> Responsibilism predicts external markers of collectivism even controlling for previous scales.                                                                 |
| 30 | <a href="#"><u>Table S4.</u></a> The US English version of the 14-item Responsibilism Scale.                                                                                                    |
|    | <a href="#"><u>Table S5.</u></a> The US English version of the 5-item Responsibilism Scale.                                                                                                     |
|    | <a href="#"><u>Table S6.</u></a> Sample demographics by culture (part 1).                                                                                                                       |
|    | <a href="#"><u>Table S7.</u></a> Sample demographics by culture (part 2).                                                                                                                       |
|    | <a href="#"><u>Table S8.</u></a> Sample demographics by culture (part 3).                                                                                                                       |
| 35 | <a href="#"><u>Table S9A.</u></a> Correlations between responsibilism and previous collectivism scales at the individual level.                                                                 |
|    | <a href="#"><u>Table S9B.</u></a> Correlations between responsibilism and previous collectivism scales at the culture level.                                                                    |
|    | <a href="#"><u>Table S9C.</u></a> Correlations between responsibilism and previous collectivism scales at the culture level after controlling for agreeableness.                                |
| 40 | <a href="#"><u>Table S10.</u></a> Divergent validity tests show responsibilism is different from Vignoles' self-construal scale components.                                                     |
|    | <a href="#"><u>Table S11A.</u></a> Correlations between responsibilism and personality.                                                                                                         |
|    | <a href="#"><u>Table S11B.</u></a> Agreeableness correlated negatively with responsibilism but positively with Singelis's collectivism scores at the individual level.                          |
| 45 | <a href="#"><u>Table S12A.</u></a> Previous collectivism scales show contradictory correlations with social status.                                                                             |
|    | <a href="#"><u>Table S12B.</u></a> Previous collectivism scales show contradictory correlations with politics.                                                                                  |
|    | <a href="#"><u>Table S13A.</u></a> Survey responses from the US and China used for comparison with                                                                                              |

researchers' guesses (part 1).

**Table S13B.** Survey responses from the US and China used for comparison with researchers' guesses (part 2).

**Table S14.** Demographics for the sample of researchers for the guessing survey.

**Table S15.** Testing whether the five-factor model applies to both individuals and cultures.

**Table S16.** Testing whether the one-factor model applies to both individuals and cultures.

**Table S17.** Testing model fit without and with an element representing response style.

**Table S18A.** Fit indices for measurement invariance across Western, "sometimes Western," and non-Western cultures.

**Table S18B.** Fit indices for measurement invariance across 100 cultures.

**Table S19A.** Culture-level scale reliability.

**Table S19B.** Exploratory factor analysis loadings.

**Table S20.** Individual-level scale reliability.

**Table S21.** Recruitment methods by culture.

**Table S22.** Responsibilism scores by culture.

**Table S23.** Sample demographics excluding non-locals and careless responders (part 1).

**Table S24.** Sample demographics excluding non-locals and careless responders (part 2).

**Table S25.** Sample demographics excluding non-locals and careless responders (part 3).

**Table S26.** Initial responsibilism scale before item selection based on validity and reliability testing.

**Table S27A.** Sources and sample sizes for previous collectivism scales used to test individual-level correlations with responsibilism.

**Table S27B.** Sample demographics for testing previous collectivism scales' culture-level predictive power before inclusion criteria (part 1).

**Table S27C.** Sample demographics for testing previous collectivism scales' culture-level predictive power before inclusion criteria (part 2).

**Table S27D.** Sample demographics for testing previous collectivism scales' culture-level predictive power after inclusion criteria (part 1).

**Table S27E.** Sample demographics for testing previous collectivism scales' culture-level predictive power after inclusion criteria (part 2).

**Table S28.** Responsibilism differences by age, gender, and social class.

**Table S29.** Cultural differences are similar before (top) and after (bottom) controlling for demographics.

**Table S30.** Spatial auto-correlation: Responsibilism predicts behavioral differences after controlling for latitude and longitude.

**Table S31.** Spatial auto-correlation: Potential ecological causes controlling for latitude and longitude.

**Table S32.** Wordings for the five most-cited collectivism scales.

**Supplementary References.**

## 1. Does Responsibilism Predict Actual Behavior?

One criticism of the “common view” is that experts agree that East Asia is collectivistic<sup>12</sup>, but they’re just wrong<sup>13</sup>. Maybe the reality is that East Asian cultures are not collectivistic. Simply showing that East Asian cultures score higher on responsibilism than Western cultures doesn’t prove that we’re actually measuring collectivism. Thus, we tested whether our culture scores for responsibilism correlated with external markers of collectivism.

We gathered external markers of collectivism from previous studies. Most of these markers are people’s behavior, although some are survey responses or vignettes. These markers tapped into responsibilities in close relationships versus strangers and anonymous systems. Extended Data Tables 1-2 report the data sources and theoretical rationales.

Responsibilism scores correlated highly with markers like extended family living together (measured in the Global Collectivism Index), family lineage systems (measured in the kinship intensity index), returning a lost wallet to strangers, and the expectation to lie for a close friend in court. [Figures S1](#) and [S2](#) graph several of these correlations. Extended Data Tables 1-2 report the raw correlations and sample sizes.

If we assume that resource scarcity is a cause of responsibilism, it would not make sense to control for economic development to test whether responsibilism is a cause or a predictor of behaviors across cultures. However, controlling for economic development could make sense to understand how much variation in cultural differences responsibilism can explain *beyond* GDP. Thus, we ran partial correlations predicting external markers of collectivism after controlling for economic development (Extended Data Tables 1-2). The results showed that responsibilism continued to predict cultural differences even after accounting for economic development. For example, responsibilism predicted a basket of Census indicators like family size and living alone before controlling for economic development,  $r(97) = 0.75$ ,  $P < 0.001$ , 95% CI [0.65, 0.83], and after,  $r(95) = 0.56$ ,  $P < 0.001$ , 95% CI [0.39, 0.68]. In sum, responsibilism is picking up on cultural differences besides economic development.

## 2. Convergent Validity and Societal Markers

We expected that responsibilism would correlate with other cultural characteristics that tap into collectivism. One metric we looked at was relational mobility<sup>20</sup>. Cultures with high relational mobility have more opportunities to make new friends and the freedom to leave relationships if they want to. Western cultures like the US tend to score high on relational mobility, while the Middle East and East Asia tend to score low<sup>20</sup>. Because relational mobility is about flexible, non-committed relationships, we expected that responsibilistic cultures should score lower on relational mobility. Extended Data Fig. 3 shows that responsibilistic cultures scored lower on relational mobility (Extended Data Table 5).

We also compared responsibilism culture scores with previous research on norm tightness around the world<sup>54</sup>. Cultures that have tight norms (such as Pakistan) report having more social restrictions on their behavior, whereas cultures with loose norms (such as New Zealand) report having less social restrictions and more freedom to behave how they like. We expected that responsibilistic cultures would tend to have tight norms.

Extended Data Fig. 4 shows that responsibilistic cultures tended to have tighter norms (Extended Data Table 5).

Extended Data Tables 3-5 report relationships between responsibilism and a list of ecological/historical factors, modern societal differences, and correlates. These markers come from a wide range of fields, such as economics, public health, and law. For example, wealthier cultures tended to be less responsibilistic. Responsibilistic cultures tended to agree less with divorce and trusting strangers. Cultures with a history of rice farming tended to be more responsibilistic. These correlations give an overall picture of how responsibilism relates to environmental and societal conditions, as well as attitudes in international surveys.

### 3. Demographic Questions and Cultural Differences Controlling for Demographics

**Demographic Questions:** We asked participants questions to understand their backgrounds. Participants reported their:

- i. Gender
- ii. Age
- iii. Marital status
- iv. Birthplace
- v. Place they primarily grew up
- vi. College major
- vii. Religion
- viii. Cultural heritage
- ix. Social status
- x. Social politics
- xi. Economic politics

We adjusted the cultural heritage and religion questions based on the local culture. For example, in Japan, we included Shintoism as a response option, but not in other cultures. We based the options on population statistics about the distribution of cultural groups and religions in each culture, as well as the recommendations of local researchers.

We followed previous research asking social and economic politics separately<sup>44,56</sup>. The questions asked, “How would you describe your political outlook with regard to social [economic] issues?” Participants rated their politics from 1 (*very liberal*) to 7 (*very conservative*).

Translating political categories across cultures is difficult. Categories that work in some cultures are less familiar in others. To lessen this problem, we worked with native speakers in each culture to choose wordings that would be understandable locally. But we recognize that a simple scale question may not do a great job of tapping into people’s political attitudes in every culture.

To measure social status, we used the ladder question, which many studies have used<sup>57</sup>. The question shows a ladder with the steps labelled from 0 to 10 and an explanation:

**Think of this ladder as representing people's social status in the country you live in.**

At the top of the ladder are the people who are the best off - those who have the most money, the most education, and the most respected jobs. At the bottom are the people who are the worst off - who have the least money, the least education, and the least respected jobs or no jobs.

**Where would you place yourself on this ladder compared to people in the country you live in?**

Participants rated their status from 0 (*bottom*) to 10 (*top*). The original item wordings (including all translations) are available on [the OSF](#).

**Cultural Differences Controlling for Demographics:** It was not feasible to collect nationally representative random samples in 100 cultures, so we aimed to collect samples with similar demographics across cultures. We did this by sampling college students wherever possible. Collaborators in a few cultures did not have access to student samples, so they used adult samples ([Table S6](#) lists sample sources). Students are in a similar stage of life, education level, and occupation, which minimizes sample differences.

However, there could still be demographic differences across cultures that influence our estimates of cultural differences in responsibilism. For example, university graduates in Greenland and some European cultures tend to be in their late 20s<sup>49</sup>.

To understand whether cultural differences were confounded by demographic differences, we ran four types of tests:

**Test 1: Estimating Culture Scores Controlling for Demographics:** First, we ran regressions with age, gender (coded as 0 = male, 1 = female), and socioeconomic status ([Table S28](#)). The results showed that older people were less responsibilistic, as were people of higher socioeconomic status.

Next, we saved the responsibilism scores adjusted for these demographics (unstandardized residuals). We used these adjusted scores to create culture scores for responsibilism. Then we compared cultural differences with and without adjusting for demographics. One culture (Singapore) did not ask the socioeconomic status question, so we ran comparisons without Singapore.

Cultural differences between non-Western and Western cultures were significant before,  $t(65) = 16.30$ ,  $P < 0.001$ ,  $d = 4.58$ , 95% CI [3.61, 5.53], and after adjusting for demographics  $t(65) = 16.41$ ,  $P < 0.001$ ,  $d = 4.61$ , 95% CI [3.64, 5.57]. This suggests that cultural differences are robust to considering demographic differences.

**Test 2: Comparing Individual-Level Cultural Differences Controlling for Demographics:** Next, we tried a different method of accounting for demographic differences. This time, we looked at cultural differences at the individual level, rather than the culture level. We ran the analysis in three steps:

- First, we limited the sample to participants who had complete data for age, gender, and socioeconomic status. This ensures that the sample is the same across different analyses.
- Then we ran a regression to test for differences between Western and non-Western cultures ([Table S29](#)).
- After that, we added demographics to the regression to see whether adding demographics made cultural differences smaller ([Table S29](#)). If so, that would suggest that cultural differences were partly confounded with demographic differences. If not, it would suggest that cultural differences were independent of demographic differences.

The results showed that cultural differences were similar before,  $\beta = 0.54$ ,  $t = 59.43$ ,  $P < 0.001$ ,  $B = 0.42$ , 95% CI [0.41, 0.44], and after accounting for demographics,  $\beta = 0.53$ ,  $t = 58.25$ ,  $P < 0.001$ ,  $B = 0.42$ , 95% CI [0.40, 0.43] ([Table S29](#)). This suggests that cultural differences were independent of differences in age, gender, or socioeconomic status.

**Test 3: Comparing Individual-Level Cultural Differences Controlling for Demographics with Random Slopes:** Finally, we tried controlling for demographic differences using random slopes across cultures in hierarchical linear models. This is different from controlling for demographics in Test 2 because it allows the effects of demographics to be different in different cultures. For example, let's assume that women are more responsibilistic than men in China, but the opposite is true in the US. Allowing random slopes allows for different gender differences in each culture. In this way, we can avoid assuming that demographic differences are the same in all cultures.

To do this, we compared two models:

**Model 1:** A hierarchical linear model that had overall effects of demographics (the same across cultures) using R code in this format:

```
lmer(Responsibilism ~ WestVsNonWest + Age + SES + Female + (1|Culture), data = triad))
```

**Model 2:** A hierarchical linear model that added random slopes (different across cultures) using R code in this format:

```
lmer(Responsibilism ~ WestVsNonWest + Age + SES + Female + (1 + Age + SES + Female|Culture), data = triad))
```

Cultural differences between Western and non-Western cultures were similar without random slopes ( $B = 0.42$ , 95% CI [0.37, 0.47],  $t = 16.26$ ,  $P < 0.001$ ) and with random slopes ( $B = 0.37$ , 95% CI [0.32, 0.42],  $t = 16.53$ ,  $P < 0.001$ ). We tested whether these estimates of cultural differences were significantly different by converting them to  $Z$  scores. This test found that the cultural differences were not significantly different  $Z =$

1.27,  $P = 0.205$ . This suggests that cultural differences were independent of demographic differences.

However, there are limitations to this analysis. These demographics are not all of the demographic differences possible between participants. In addition, the socioeconomic status question asks people to compare themselves to other people in their same country. This has advantages, but it is also only one measure of socioeconomic status. For example, it may also be useful to measure differences like family income and parental education separately. Thus, we are confident that cultural differences are independent of basic demographic confounds, but, as with any study, we cannot rule out confounds in demographic differences that we did not measure.

**Test 4: Limiting the Sample to Participants Under 24 Years Old:** We ran an analysis of cultural differences after limiting the sample to people under 24 years old. We did this to deal with the fact that university students in some education systems are older than others<sup>49</sup>. We chose this age range because it would cover most students in education systems that have a structure like the US, making the sample age range more similar across cultures.

Like in Test 2 above, we ran regressions comparing West vs. non-West cultural differences before and after limiting the sample by age. Results showed that cultural differences were similar with the full age range and with the limited age range ([Table S29](#)). If anything, cultural differences were significant after restricting the sample ( $\beta = 0.56$ ,  $P < 0.001$ ,  $B = 0.44$ , 95% CI [0.42, 0.45]) and with the full age range ( $\beta = 0.54$ ,  $P < 0.001$ ,  $B = 0.42$ , 95% CI [0.41, 0.44]). This suggests that cultural differences are robust to restricting the age range.

#### 4. Sample Selection

For each culture, we set a minimum goal of 50 participants who were born and raised in the culture. Fifty participants is a small sample. We aimed to collect larger samples in many cultures. However, we identified 50 as a minimally feasible goal in cultures where it is not common to run psychology studies.

In [Supplementary Section 34](#), we report analyses with random sub-samples ranging from 5 participants to 100 participants per culture. These analyses let us estimate how robust the findings are to samples of different sizes. For example, random sub-samples of just 40 participants per culture produced highly similar external validity correlations as the full sample ([Fig. S18A](#)). This gives some confidence that samples of this size can produce reliable results.

Another way to judge sample sizes is with the effect size. The large effect size between Western and non-Western cultures (culture level:  $d = 4.58$ , individual level:  $d = 1.35$ ) means that a sample of 50 people in two groups still gives 99% statistical power. Despite the statistical power and the robustness checks of the random sub-samples, we acknowledge that a limitation of our sample is that the samples for some cultures are small. Larger samples would provide more precise estimates of cultural differences.

Our goal was to sample college students because (a) these samples are easier to reach across many cultures and (b) comparing college students makes the samples comparable in age, education, social status within the country, and income. We also recruited participants through the Prolific platform. Prolific allowed us to selectively recruit



participants who were current students and born and raised in the culture. In addition, we recruited college students in some cultures through the survey company Cint. [Table S21](#) lists the recruitment method for each culture.

## 5. Scale Development and Translation

**Scale Development:** The first author was primarily responsible for developing the scale items. He first developed the items by surveying previous collectivism scales and identifying whether they fit with the conceptualization of responsibility or not. He identified three items from previous research that mentioned responsibilities in close relationships:

- i. We should keep our aging parents with us at home.
- ii. Children should feel honored if their parents receive a distinguished award.
- iii. Children should be taught to place duty before pleasure.

Surveying previous scales was also valuable in giving an understanding of what changes would be useful in creating scale items. [Table S32](#) lists the scale items for previous scales. For example, looking at previous scale items, many used vague relationship terms like “others” or “group.” Even phrases like “close relationships” that specify closeness still may be a bit abstract or vague. Many items talked about feelings (such as “feel hurt”) or more abstract ideas like identifying with people (such as “my close relationships are an important part of my self-image”). Some scale items used terms that probably feel abstract or academic in many communities or languages, such as “self-interest” and “work-related matters.” It was helpful to identify these wordings that we hypothesized were straying from responsibilities in close relationships (and perhaps straying from guidelines about effective scale design for studies across cultures<sup>10</sup>). Finding these wordings helped guide decisions to use more concrete language with specific situations.

After identifying items in previous scales that fit the concept of responsibility, the first author created more items, for a total of 26 items ([Table S26](#)). To create the items, he followed the recommendation from Peng and colleagues to use concrete behaviors in specific situations, rather than vague, abstract wordings<sup>10</sup>. Another goal was to sample a broad range of relationships, from clearly close relationships to clearly distant relationships (such as strangers). He also made effort to sample relationships that are not as close as family or friends but still belong to the same larger group, such as an unfamiliar classmate and coworker. The goal here was to test the idea that belonging to the same large group is not sufficient for people to feel duties in responsibility cultures.

The first author also used his background in cultural psychology, experience living in China, and Chinese language ability to help do a first-pass check on whether the items would seem natural to participants in China. He worked with a native speaker research assistant from China (then living in the US) to cross-check his understanding on whether items would make sense to participants in China. Supplementary Section 26 describes the bilingual scale creation process in more detail.

“Natural” includes understanding the words, using words that would be common in everyday life (for example, not “authority figures” or “detested”), and hopefully easy for people to imagine themselves doing. This was the goal, although the translation process



into other languages and scale refinement tests eventually identified shortcomings of some items. In other words, the first attempt was not successful for every item.

Of course, English and Chinese are not representative of the whole world's languages and cultures. When translating the scales for use in other cultures, we gathered input from professional translators and researchers. The next section describes some of the changes that feedback generated.

In creating the scale, an intentional decision was to include items that people in responsibilistic cultures would agree with and items that people in individualistic cultures would agree with. This is important (a) to capture a broad range of items and (b) to allow us to differentiate between people who simply tend to agree with all items and actual differences in attitudes. This decision was based on previous research finding that cultures differ in their response styles, which can add noise to cultural comparisons<sup>17,58,59</sup>.

In the early rounds, we ran the scale with participants in the US and China and solicited their feedback, such as whether there were any confusing parts or wordings that were difficult to understand. This led us to make changes to the formatting of the vignettes to make the process clearer. For example, we added a line break in Item 4 between the main vignette and the prompt, "I would feel comfortable telling this co-worker about the account."

Different items have different formats. Some are simple statements, and some are longer vignettes that have a simple statement after the vignette. The main reason for the difference is that we had two sources of items. Some of the items came from scanning previous scales for items that matched the concept of responsibilism, and those items are all short. In contrast, the other source was our newly created items, which are longer. In creating our items, we wanted to follow the advice of Peng and colleagues<sup>10</sup>, who found that vignettes worked better than more abstract statements. That led us to create longer scenarios, which we hypothesized would have the upside of being better at producing valid results, but with the downside is that their format differs from the standard items.

In each new country and language we rolled out the survey in, we solicited feedback from people in that country. This included professional researchers and translators. In Supplementary Section 5 and Table S26, we describe some of the edits we made to the scale based on feedback from people across cultures.

**Translation:** After developing the scale, we translated the scale into local languages with the help of professional translators and local researchers. In most cases, we first hired a professional translator to translate the scale. Then we asked the local researcher to check the translation and search for any parts that might be hard to understand.

Our goal was to maximize the equivalence of the intended meaning of the items across cultures, rather than strict word-for-word translations. When questions arose about how to translate questions, we explained the intended meaning of the original items and encouraged the local researchers to translate the items so that they conveyed the same meaning.

In a few cases, it was difficult to literally translate the items because the original reference was uncommon in the local culture. For example, one item mentions drinking beer. Several researchers in majority-Muslim cultures suggested that these would be

inappropriate in their cultures. They suggested changing the items to be about drinking tea (which is common in public), and we followed these suggestions.

Another item mentions flying on a plane. In some developing countries, local researchers suggested that many locals would not have experience on planes. We followed their suggestion to change this to a long-distance bus ride.

Finally, another item mentioned college students dating. In a few cultures, local researchers suggested that this would be uncommon. The question asked about close friends giving harsh feedback about the relationship. We decided to change this to be about friends giving harsh feedback about the respondent's choice of career.

The results of the measurement equivalence tests suggest that participants understood the items similarly across cultures ([Supplementary Section 10](#)). We also tested people across cultures in English in a few cases where English is common across cultures, such as comparing the US to Hong Kong, Singapore, and South Africa. The fact that the cultural differences persisted, even when the items were in the same language, suggests that cultural differences are not a confound of translation.

However, we encourage future researchers to continue to test refined versions of the scale to improve its equivalence and generalizability. Qualitative research may also help further refine the items. In-depth conversations with people across cultures could help generate new items that might prove useful to distinguish facets of responsibilism.

## 6. Scale Refinement and Item Selection

We started the project with 26 items to measure responsibilism and ended up with 14 items after testing the scale across cultures. We refined the scale by removing poorly performing questions. In this section, we detail the process of refining the scale items. But before we do that, we first test whether the main conclusion of the paper remains true before the process of item selection.

**Cultural Differences Are Significant Before Any Scale Refinement:** As a test of our theory that responsibilism is a better description of collectivism than the classic collectivism scales, we think it is important to test whether the full scale (before item selection) could detect cultural differences. If not, it could suggest that we are merely cherry picking the items that show cultural differences and the general theory about responsibilism is wrong. Therefore, we tested differences in total responsibilism scores across all of the original items between West versus non-Western cultures ([Supplementary Section 24](#) defines these categories).

Before any item selection, non-Western cultures scored significantly higher on responsibilism than Western cultures,  $t(66) = 10.96$ ,  $P < 0.001$ ,  $d = 3.07$ , 95% CI [2.31, 3.82]. We also compared the West with East Asia because this is the more specific claim of cultural differences in the research on collectivism in cultural psychology. East Asian cultures scored higher on responsibilism than Western cultures,  $t(22) = 9.11$ ,  $P < 0.001$ ,  $d = 4.09$ , 95% CI [2.59, 5.56].

In sum, the responsibilism items continued to show cultural differences even before any item selection. This suggests that the top-down theory about responsibilism was correct. In other words, if we merely rely on having what we think is the correct theory of what collectivism is, then we can create a scale that will reliably detect the cultural differences that earlier collectivism scales failed to find. Put another way, the results from

this section are inconsistent with the idea that these cultural differences are an artifact of selecting only the items that work.

**Item-Total Correlations:** Our main process for item selection was to calculate item-total correlations. These correlations ask how well each question correlates with overall  
5  
responsibilism scores. If an individual question correlates negatively with overall  
responsibilism scores, it suggests that the question is not working as we expected it to.

We documented the step-by-step process of item selection with analysis syntax and  
10  
results. We started with the culture-level dataset and looked for items with negative item-  
total correlations. We deleted the item with the worst correlation and then calculated  
item-total correlations again. We did this until we removed all items with negative item-  
total correlations in the culture-level dataset. This analysis identified 10 items to remove.

Next, we continued the tests with the individual-level dataset. This analysis identified  
15  
one item to remove. Finally, the two main authors discussed the remaining items. We  
decided to remove item 25 because we thought most people around the world would have  
a hard time relating to being a judge ([Table S26](#)). [Table S4](#) lists the final scale items.  
[Supplementary Section 11](#) reports the reliability.

**Checks on External Validity Before and After Scale Refinement:** We tested  
20  
whether the process of selecting items improved tests of external validity. We did this by  
testing how well behavioral markers of collectivism around the world correlated with the  
full original responsibilism scale versus the refined scale. The external markers were  
seven measures, such as kinship practices, family practices, and lying to protect a friend.  
([Supplementary Section 1](#) describes the measures in detail.)

Before refining the scale, the responsibilism culture scores had an average absolute  
25  
correlation with the external markers of  $r = 0.546$ . After refining the scale, the correlation  
increased to  $r = 0.661$ . This suggests that the process of refining the scale led to a more  
accurate measure of collectivism.

We also tested whether refining the items made the scale more sensitive to picking up  
30  
cultural differences. To do this, we tested for differences in responsibilism between  
Western and non-Western cultures. Refining the scale increased the sensitivity to cultural  
differences from  $t(66) = 10.96$ ,  $P < 0.001$ ,  $d = 3.07$ , 95% CI [2.31, 3.82], to  $t(66) = 16.34$ ,  
 $P < 0.001$ ,  $d = 4.58$ , 95% CI [3.62, 5.52].

This result suggests that refining the scale items made the scale more sensitive to  
35  
measuring cultural differences. Internal validity is not always required for external  
validity<sup>60</sup>. However, the analyses found that improving the internal validity of the  
responsibilism scale improved external validity.

## 7. Component Scores and Factor Structure

We ran factor analyses to test whether it makes more sense to think of responsibilism  
40  
as a single factor or having multiple factors. We started by running exploratory factor  
analyses and by looking for commonality among the items that tended to correlate with  
each other. With each model, we tried splitting out more factors or collapsing into fewer  
factors to see if model fit improved.

The final results showed good fit for both a five-factor model and a single-factor  
model ([Tables S16-S17](#)). This result is similar to personality facets. For example, the

extraversion facet of the Big Five scale has six sub-components: warmth, gregariousness, assertiveness, activity, excitement-seeking, and positive emotions. With personality, it can be useful to talk about the facet of extraversion or to break extraversion down into sub-components. We think it makes sense to think of collectivism similar to personality facets in that both can be collapsed into a larger component or split out into more specific sub-components.

Should researchers use the overall scores or the sub-factors? We think researchers will find different scores useful depending on what their research question is. Like with extraversion and the Big Five, we suspect most researchers will be more interested in the overall scores than the sub-facets. However, researchers specifically interested in family relationships might find more predictive power in the family responsibilities sub-facet. Researchers interested in friend relationships might get more usefulness out of the tight friendship sub-facet. We provide scores on the sub-facets for all cultures on [the OSF](#).

Maps for the component scores are in [Figures S6-S10](#) and [online](#). We titled the five sub-components:

- i. **Family responsibilities:** These items have to do with teaching children duties, ties within the family, and taking care of family members. The items are 1, 6, and 8 in [Table S4](#).
- ii. **Guarding personal details:** These items have to do with being cautious about sharing information with non-close acquaintances or strangers. We say “non-close acquaintances” because two of the items ask about people who are in the same large group (a company) but that the participant does not know personally. The items are 4, 5, and 10 in [Table S4](#).
- iii. **Tight friendship:** These items have to do with sticking with friends, even if they do things that make you unhappy. One item asks about making new friends with strangers. These items are similar to low relational mobility<sup>20</sup>, although the correlation between tight friendship and relational mobility suggests there is some room for difference between the two:  $r(53) = -0.43$ ,  $P = 0.001$ , 95% CI [-0.62, -0.19]. Furthermore, relational mobility had similar correlations with family responsibilities  $r(53) = -0.51$ ,  $P < 0.001$ , 95% CI [-0.68, -0.29] and fitting in  $r(53) = -0.52$ ,  $P < 0.001$ , 95% CI [-0.69, -0.29]. The items are 9, 11, and 13 in [Table S4](#).
- iv. **Stranger danger:** These items have to do with being wary of sharing information with or trusting strangers. The logic is that, in cultures where people have tight ties and responsibilities in close relationships, people may have fewer responsibilities for strangers and also less trust. This fits with the finding that cultures that agreed more with responsibilities in the family also tended to report more stranger danger,  $r(98) = 0.46$ ,  $P < 0.001$ , 95% CI [0.61, 0.29]. The items are 2a, 2b, and 3 in [Table S4](#).
- v. **Fitting in:** These items have to do with going along with what other people are doing and avoid standing out. Another way to look at the items is people who value fitting in tend not to be bold, direct communicators. For example,

one item asks about approaching a classmate you've never met before without an introduction. We wrote the classmate question based on the idea that, in responsibilistic cultures, the unknown classmate in a large lecture class would not count as a close relationship. Talking to the unknown classmate would not be fitting in to social norms because (1) breaking the ice requires boldness and (2) relationships come tied up with responsibilities, so it is not adaptive to cast a broad "relationship net." The items are 7, 12, and 14 in [Table S4](#).

With the "fitting in factor," we also asked whether it might be measuring hierarchy because two of the items mention a boss. We explored this possibility by testing correlations between the "fitting in" facet and power distance scores from the GLOBE study<sup>61</sup>. When testing this question, it is important to keep in mind that collectivistic cultures tend to score higher on power distance<sup>61,62</sup>. For example, the Global Collectivism Index is correlated with power distance  $r(48) = 0.46$ ,  $P < 0.001$ , 95% CI [0.21, 0.65]. Thus, we should ask whether the "fitting in" factor had a particularly high correlation with power distance or just a moderate correlation that we would expect from the general correlation between collectivism and power distance.

"Fitting in" had a modest correlation with power distance practices  $r(48) = 0.34$ ,  $P = 0.015$ , 95% CI [0.07, 0.57]. Correlations were similar with the other factors (such as stranger danger,  $r = 0.31$ ,  $P = 0.026$ , 95% CI [0.04, 0.55]). These results suggest the "fitting in" facet does not uniquely tap into hierarchy.

Another way to test this question is to see if "fitting in" scores are correlated with actual inequality in societies. We tested this using the Gini index across cultures (Extended Data Table 5 lists the data source). However, "fitting in" was not significantly correlated with inequality,  $r(64) = 0.10$ ,  $P = 0.432$ , 95% CI [0.15, 0.33]. Thus, the results suggest that fitting in is not strongly related to hierarchy.

If "fitting in" is not particularly related to hierarchy, what is it related to? We found some evidence that it's related to self-expression and self-esteem. Self-expression and self-esteem make sense because bucking social norms should be easier for people who have high self-esteem or who value self-expression.

"Fitting in" was negatively correlated with self-expression across cultures from Vignoles and colleagues' study<sup>45</sup>,  $r(29) = -0.51$ ,  $P = 0.003$ , 95% CI [-0.73, -0.19]. "Fitting in" was also negatively correlated with self-esteem values across cultures from Schwartz<sup>63</sup>,  $r(44) = -0.41$ ,  $P = 0.004$ , 95% CI [-0.63, -0.14]. In sum, "fitting in" seems to be tapping into the conflict between fitting in with social norms versus being bold, expressive, and a direct communicator.

We should be clear: Our goal in the study was to find a reliable measure of collectivism across cultures. We created these items based on an understanding of responsibilism and recent research on tight social ties in collectivistic cultures<sup>20,27,64,65</sup>. However, we did not exhaustively test all possible components of collectivism. There could easily be components that we overlooked. Thus, we do not think this version of responsibilism is the absolute end of the road for measuring collectivism.

Future research can improve on the responsibilism scale. We think a worthwhile future project would be to create a broad range of potential components of collectivism

and test them extensively across many cultures. The goal would be to see which sub-components of collectivism tend to occur together.

Ideally, this research would also work hard to stress test those items. For example, the research should take into account correlations with markers of external validity (like the markers we use in this study). The research should also try to rule out misleading correlations. For example, some items might correlate with collectivism because collectivistic cultures tend to be poorer. This probably explains why Hofstede's classic individualism scale includes an item about the importance of "good ventilation and lighting" at work. ([Supplementary Section 27](#) explains this in more detail.)

In short, this responsibilism effort is not the end of the road. Just like the Big Five personality scale was the result of decades of work from different labs, more work needs to be done to refine the concept of responsibilism. We hope that we can kickstart this process by opening up this big, global dataset to other researchers.

## 8. Responsibilism Varies Significantly Between Cultures

One question that comes up in cultural research is whether there is meaningful variation between cultures compared to variation between individuals. As a silly example, if we measure flying the American flag around the world, culture would explain a large percentage of the variation around the world. Americans would fly the flag a lot, and people in other cultures would fly it very little. In this case, culture would explain a lot of variation.

As another silly example, we could measure whether people have a Y chromosome. Around the world, about half of people (men) will have a Y chromosome. There will be small variations across cultures due to factors like war (which tends to kill men) and selective abortions (which are more common against female children in China). But overall, there will be a lot of variation between individuals and very little variation between cultures. In that case, using culture to analyze Y chromosomes would have far less value than using culture to analyze flag use.

We tested this by estimating the percentage of variation in responsibilism between cultures versus between individuals within cultures. To do this, we calculated the intra-class correlation coefficient from a hierarchical linear model using the LMER function in the program R. Culture explained 30.6% of variation.

In sum, there was a meaningful amount of variation between groups. This suggests it is meaningful to analyze differences based on cultures. However, this does not rule out meaningful differences between regions within national cultures (such as our comparison of the US South and North) and individuals (such as our comparison of social classes).

## 9. Responsibilism Can Measure Individual Differences and Cultural Differences (Cross-Level Measurement Invariance)

In measuring cultural differences, it is important to understand the difference between individual differences and cultural differences. Individuals and cultures are not always the same. For example, in the US, individual income predicts voting for conservatives. But state-level average income predicts voting for liberals<sup>66</sup>.

Similarly, some concepts hang together at the group level but not the individual level (or vice versa). One nice example is hierarchy<sup>63</sup>. When measuring cultures, hierarchical



5 cultures tend to have more submissiveness (among subordinates) and more dominance (among leaders). So if we're comparing cultures, submissiveness and dominance will be correlated. But if we measure individuals, submissiveness and dominance will probably be negatively correlated because submissiveness and dominance apply to people who are at different places in the hierarchy.

For this reason, we have to check whether our concepts are actually similar at the culture level and the individual level before we make comparisons at both of these levels. For example, in this paper, we use responsibilism to compare individuals based on social class. Is this justified?

10 To be clear, measures don't need to be valid at both levels. It is possible to have a valid measure for cultures that simply doesn't apply for individuals. It is possible to have a valid measure for individuals that doesn't work for cultures. There's some evidence that personality scales don't work well to measure cultural differences<sup>67</sup>.

15 We tested this by running models of cross-level measurement equivalence. We tested both the one-factor model and the five-factor model because the factor analysis suggested that both of these models fit the data well. These measurement equivalence models ask whether the same factor structure applies to both individuals and cultures. For example, with the hierarchy example, if assertiveness and submissiveness are positively correlated at the culture level but negatively correlated at the individual level, this test would  
20 conclude that the factor structure is not the same for both levels.

The results suggest that responsibilism has a similar factor structure at the individual and culture levels ([Table S15-S16](#)). In the tables, Model 4 tests how well the model fits if we force the model to have the same structure for both individuals and cultures.

25 We can compare results two ways. (a) We can look at the *absolute* measures of model fit in isolation, and (b) we can compare the *relative* measures of model fit in the final model to the model before it, which does not force the same factor structure for both levels. If the model fit changes for the worse after forcing the structure to be the same at both levels, it suggests that it should not be the same. The notes below each table report suggested cutoffs for judging these metrics of model fit.

30 Both ways of looking at the results suggest similar conclusions: the model fits well at the individual and group level. Looking at the absolute measures of model fit for the final model, the CFI and RMSEA satisfied the recommended cutoffs. Comparing the *relative* fit of the final model (Model 4) to the model before it (Model 3) also satisfied recommended cutoffs for both models. Thus, the metrics of model fit generally supported  
35 the idea that the one-factor and five-factor models apply to both individuals and cultures.

40 However, there is one exception worth noting: the results for SRMR satisfied the cutoffs for the individual level, but not the culture level. The SRMR for the five-factor model was slightly above the cutoff for the five-factor model (0.10 vs. cutoff of < 0.08) and even higher for the one-factor model (0.32). Thus, there it might be possible to improve the model at the culture level. Or we might get better model fit if we use a slightly different model for the culture level than the individual level.

However, the results for the BIC measure of model fit led to the opposite conclusion. The BIC showed the best model fit for the final equivalent model in both the one-factor model and the five-factor model. Thus, although not every metric led to the same



conclusion, many of the metrics supported the idea the same factor structure applies to both individuals and cultures. These results mostly support the idea that we can use the responsibility scores to compare individuals and cultures.

## 10. Scale Equivalence Across Cultures: Measurement Invariance

We ran analyses of cross-group measurement invariance to test whether people interpreted the responsibility scale similarly across cultures. We ran cross-level models in the program MPlus (Version 8). These models account for both individuals and cultures at the same time.

We ran two types of invariance tests:

**i. Groups of Cultures:** In one analysis, we ran tests of cultures grouped into Western, “sometimes Western,” and non-Western cultures. This tests for invariance across broad cultural groups.

**ii. 100 Cultures:** In the other analysis, we ran tests of participants grouped into all 100 cultures.

### Test i: Groups of Cultures

Researchers have suggested that structural equation models need at least 200 people per group in order to have enough statistical power (44). Perhaps because many of our cultures had fewer than 200 participants, the invariance tests failed to converge. To deal with this, we ran tests while grouping cultures into Western, sometimes Western (such as Mexico), and non-Western ([Supplementary Section 24](#) defines “sometimes Western” in detail). This has the downside of lumping cultures together. However, it allowed models to converge, and the results from the main study found that this three-category grouping corresponds to large cultural differences.

The three-category grouping captures a large portion of the variation across cultures. For example, splitting out all 100 cultures explained 32.2% of variation in responsibility, but splitting cultures into just these three groups explained 24.9% of variation (calculated from one-way ANOVAs). In other words, placing cultures in these three groupings capture more than half of the variation between cultures.

**Configural invariance:** We tested for “configural invariance” in MPlus using multi-level analyses that consider both individuals and cultures. If a scale shows configural invariance, it suggests that the factor structure is the same across groups. For example, if responsibility is related to tight friendship ties in some cultures but not others, the results would show poor fit for configural invariance. As another example, in the five-factor model, if the item about aging parents fit with the family responsibilities factor in most cultures, but then it fit with the tight friendship factor in another culture, this would show poor fit for configural invariance.

We estimated a latent factor for responsibility and a common-method factor to take into account differences in response styles<sup>68</sup>. Model fit was assessed using multiple goodness-of-fit indices: comparative fit index ( $> 0.90$ )<sup>69</sup>, root mean squared error of approximation (RMSEA  $< 0.08$ )<sup>70</sup>, and standardized root mean squared residual (SRMR  $< 0.08$ )<sup>69</sup>. We did not examine chi-squared tests because they are very likely to be significant with large samples, but we report them nonetheless<sup>71</sup>.

The results showed successful configural invariance across cultures. Testing a model with cross-level invariance showed good scores on fit indices ([Table S18A](#)). The CFI was 0.93, which is above a recommended cutoff of  $> 0.90$  for acceptable fit but below another recommendation of 0.95<sup>69</sup>. The RMSEA was 0.03, below a recommended cutoff of  $< 0.08$ <sup>70</sup>. The SRMR was 0.02, below a recommended cutoff of  $< 0.08$ <sup>70</sup>. Thus, the model showed a similar structure across cultures.

**Metric invariance:** Next, we tested for metric invariance. Metric invariance starts with configural invariance and adds another layer on top. Metric invariance tests whether different cultures have similar *factor loadings* (the relationships between the individual scale questions and the overall responsibility scores). For example, if the “aging parents” question is strongly related to responsibility in China but not in Sweden, then metric invariance would show poor results.

Put another way, successful invariance suggests that the items of overall responsibility scores are similar indicators in one culture as they are in other cultures. If an item is a strong indicator in one culture, it is also a strong indicator in other cultures. The same is true of the loadings between the five components and the overall responsibility scores.

If the results support metric invariance, it suggests that people are understanding the items similarly across cultures. For example, with the responsibility scale, people seem to have similar understandings of what it means to trust strangers versus take care of elderly parents. Metric invariance also suggests that the scale was translated effectively across cultures. Sloppy translation would change how people understand the items across cultures, which would make it hard to achieve metric invariance.

The results showed successful metric invariance across cultures ([Table S18](#)). To judge metric invariance, researchers suggest comparing the *changes* in model fit moving from the configural model to the metric model<sup>72</sup>. If the model shows only small changes after adding the assumptions of invariance across cultures, then the scale fits well across cultures. If the model fit gets a lot worse, it suggests that factor loadings vary across cultures.

To judge the results, researchers have suggested cutoffs for changes to model fit compared to the model before it ( $\Delta$ RMSEA,  $\Delta$ CFI,  $\Delta$ SRMR). However, many of these cut-offs are based on multi-group and longitudinal tests among small number of groups and time points<sup>73</sup>. These cutoffs may be too stringent for constructs that are truly invariant. Researchers have found that, with more and more groups, changes in model fit grow larger (in other words, non-invariance metrics grow as the number of groups grows)<sup>73</sup>. Some researchers have suggested dealing with this problem by using cutoffs that are more appropriate for large samples and by estimating the effect size of the non-invariance, rather than just a “yes or no” cutoff<sup>74</sup>.

For metric invariance, model comparisons met even the conservative cut-off criteria recommended by past research. The change in CFI was 0.020, which equals the recommended cutoff of  $< 0.02$  specific to models with large groups<sup>73</sup>. The change in RMSEA was 0.001, below a suggested cutoff of  $< 0.015$ <sup>72</sup>. The change in SRMR was 0.005, below a suggested cutoff of  $< 0.03$ <sup>72</sup>. The results satisfied all of the cutoffs for change in model fit ([Table S18](#)). In short, the results support the idea that the

responsibilism models fit similarly across cultures. This suggests that it is valid to use responsibilism scores across cultures.

**Scalar Invariance:** Some studies include a third step of testing scalar invariance. Doing so extends metric invariance to add another layer of invariance. Scalar invariance constrains the item intercepts to be equal across groups. Constraining the intercepts effectively means that any two cultures that have the same overall responsibilism score should also have the same score for each scale item. It is essentially asking:

If two cultures have the same overall scores (latent scores) on responsibilism, do they have the same scores for this particular scale item?

For example, if China and Kuwait both scored the same on responsibilism overall, do they have the same average agreement with the item, “We should keep our aging parents with us at home?” If they do not have the same agreement with that item, the test for scalar invariance would fail.

With strict scalar invariance, it suggests that variation in responsibilism (high or low scores) causes variation in the average scores for each item (the item intercepts). If item intercepts are not equal, even after assuming that the underlying latent factor is equivalent across groups, then item intercepts may be capturing variation in characteristics other than responsibilism.

However, scalar invariance testing is controversial. On the one hand, two researchers wrote that, “Scalar non-invariance precludes any observed mean comparisons”<sup>75</sup>. Two researchers summed up what this conclusion sounds like to many researchers: “your mean differences do not mean anything. So, you are, in effect, prohibited to look at them.” On the other hand, some researchers disagree that scalar invariance is necessary to make valid comparisons across cultures<sup>60,76</sup>.

We erred on the side of providing as many tests of the responsibilism scale as possible, so we decided to test scalar invariance. We evaluated changes in model fit using criteria for large groups<sup>73</sup>. [Table S18](#) reports tests of scalar invariance.

There was some disagreement between the different model fit indicators about whether the responsibilism scale demonstrated scalar invariance. The  $\Delta RMSEA = 0.008$  and  $\Delta SRMR = 0.007$  satisfied even the most conservative cut-offs for establishing scalar invariance. These indicators suggest that the scale has scalar invariance (also called “strong invariance”) and that the item intercepts are equivalent across cultures. However, the  $\Delta CFI$  was 0.068, which is larger than the suggested cutoff. This did not meet the cutoffs, even for larger groups.

It is worth noting that scalar invariance does not have to have a dichotomous “yes or no” result. Instead, we can use effect sizes to quantify how large the departure from invariance is<sup>74</sup>. In this case, using the  $\Delta RMSEA$  and  $\Delta SRMR$  criteria, non-invariance is negligible and near-zero. For  $\Delta CFI$ , departures from invariance were very small in terms of effect size. This suggests that, if one were to choose this as the criteria and ignore the rest, the scale could reasonably be considered as invariant even though this criterion did not match what is typically preferred when employing the dichotomous cut-offs. In sum, comparative model fit suggests that scalar invariance in the responsibilism scale is also likely in our study.

However, given that the responsibilism scale did not fit the traditional cutoffs of scalar invariance, we thought it would be useful to look into *why* a scale could fail scalar invariance (as it seems that most scales do). There are different possibilities. Scalar invariance results cannot tell us whether a scale “failed” because (a) collectivism legitimately expresses itself differently across cultures (for example, maybe collectivism manifests more through the family in China and more through friends in Kuwait), (b) there are differences in response styles across cultures causing variation in the scores, (c) poor translation is causing false cultural differences, or (d) something else.

We have data that help us test possibilities b and c. We think poor translation is unlikely because of the efforts we took to create concrete scenarios that would be easy to translate. We also created the scale in English and Chinese simultaneously, which helps limit colloquial phrasings specific to English ([Supplementary Section 26](#)). The strongest evidence is the fact that cultural differences remained when comparing different cultures using the same language and when comparing Americans with different cultural backgrounds all taking the scale in English (reported in the main text). Thus, we think it is unlikely that large translation errors are causing variation across cultures.

Next, it is important to take seriously the possibility of response style differences across cultures. We accounted for response style in different ways. In a simple way, we adjusted participants’ responses by their average agreement across items (described in [Supplementary Section 18](#)). Although this method is crude, the results showed that it improved the ability of scores to predict external markers of collectivism. The fact that correcting for response style improved external validity suggests that there are response style differences across cultures. What’s more, correcting for them in this way helped improve the accuracy of the estimates of responsibilism across cultures.

A more sophisticated method is to model response style in structural equation models. We did this by testing models that added a common-method factor. We ran models with response style that was identical for every scale item (setting the loading to 1 for each item) and models that allowed response style to be different for different items (called “bifactor”). If cultural differences remain similar after accounting for response style, it suggests that response style differences are not strongly affecting the data. Culture scores that take into account response style using structural equation models were highly similar to results using the simpler method,  $r(98) = 0.94$ ,  $P < 0.001$ , 95% CI [0.91, 0.96]. Both methods increased correlations with external markers of collectivism, but the simpler method increased the correlations more ([Supplementary Section 18](#)).

These analyses help address the question of whether the responsibilism scale failed scalar invariance because of response style differences, translation, or because there are different ways to reach the same overall responsibilism score. Two methods of taking into account response style led to largely similar estimates of culture scores. Like with most scales in use in psychology today, we cannot completely rule out the possibility of response style biases between groups. However, the results of the response style analyses and the correlations with external markers of collectivism make it less likely that response style bias is a serious problem with the scales.

## Test ii: 100 Cultures

To deal with the shortcomings of invariance testing with a large number of groups, we ran multilevel confirmatory factor analysis measurement invariance testing. Researchers recommended using this approach when the number of groups is larger than ten<sup>60,77</sup>, and it is optimal when the number of groups is higher than thirty<sup>78</sup>. In other words, this method is better designed for testing invariance with a large number of groups. Using this method, we tested invariance with all 100 cultural groups ([Table S18B](#)).

We followed the procedure suggested by Jak and colleagues<sup>79</sup> and Kim and colleagues<sup>80</sup>. Jak and colleagues<sup>79</sup> explained that the cross-level invariance of factor loadings is equivalent to factor loading invariance across groups, which matches the definition of the metric measurement invariance model. The scalar model adds a constraint of zero residual variance at the culture level, which tests the idea that the intercepts are invariant across groups.

**Configural:** The configural invariance results passed all suggested cutoffs ([Table S18B](#)).

**Metric:** The metric invariance results passed the cutoffs for RMSEA and SRMR. The change in CFI (0.03) was higher than the recommended cutoff (0.02). However, readers may consider the size of the deviation (0.01) and the fact that the absolute value of the CFI remained above 0.90.

**Scalar:** Researchers have argued that scalar invariance is an “unachievable ideal that in practice can only be approximated”<sup>81</sup>. Despite that high bar, the scalar invariance results passed the suggested cutoffs for RMSEA and SRMR, but not CFI. Given that some researchers have argued that scalar invariance is an unachievable ideal, we interpret these results as fairly positive, especially with a large number of groups.

**Invariance Summary:** In sum, the results of the invariance analysis suggest that the responsibility scale has a similar structure across cultures. Responsibility seems to be something that people can report on similarly across cultures. Finally, the results suggest that the factor structure is similar both for individuals and for groups.

Although we tested for invariance, we also take seriously the arguments of researchers who say that we should avoid the trap of narrowly focusing on fit metrics. Some researchers have argued that external validity is more important—does this predict actual behavior<sup>60</sup>?

Other researchers have argued that it is unreasonable to expect all scales to meet these cutoffs and that we should instead use these metrics to compare progress over previous results<sup>45,82</sup>. For example, researchers have found that the most-cited collectivism scales score poorly on metrics like the CFI. While the CFI for responsibility was above 0.90 in most models, the CFIs for the Singelis 1994 scale range from 0.25 to 0.65, and the CFIs for Gudykunst and colleagues’ 1996 scale range from 0.55 to 0.72 (see the discussion in Vignoles and colleagues’ paper)<sup>45</sup>. Compared to these scales, the responsibility scale seems to be an improvement.

Because of the critiques of invariance testing, we place more emphasis on tests of external validity in this study. We ask whether the responsibility scale correlates with external markers of collectivism, including several markers of behavior, rather than self-report. We also ask whether the responsibility scale can pick up on differences based on

social class because behavioral studies have found collectivism differences based on social class<sup>44,56</sup>. Invariance testing may be helpful, but it should be interpreted along with other important tests of scale performance.

## 11. Scale Reliability

Next, we tested the factor structure and reliability of the responsibility items. We tested factor loadings in the five-factor structure described in [Supplementary Section 7](#). To test it, we ran a multigroup factor analysis in MPlus. The analysis takes into account that individuals are nested in cultures.

The analysis has an overarching responsibility factor with five factors under it. Each factor then has the individual survey questions under them. The model includes a factor to represent response style (often called a “common-method factor”). This factor helps account for response style differences across cultures, such as whether people in some cultures tend to agree more in general.

[Fig. S16](#) shows the factor loadings. The note to the figure reports the fit statistics. The fit statistics were acceptable compared to suggested cutoffs. We originally ran a model with freely estimated loadings between responsibility and the five factors. However, the model failed to produce fit statistics. Thus, we constrained factors three and five to be equivalent to each other because they are correlated at the cultural level, and the model produced fit statistics.

We also ran tests of weighted reliability (also called “coefficient h”) and omega as measures of scale reliability ([Tables S19-S20](#)). Weighted reliability gives different weights to different scale items, similar to how structural equation models give different weights to different items (such as [Fig. S16](#)). If items in a scale correlate with each other more strongly, then reliability scores will be higher. Researchers have suggested cutoffs of > 0.50 or 0.60 for scale reliability<sup>83,84</sup>. At the culture level, weighted reliability was 0.89, and omega was 0.83 ([Table S19](#)). At the individual level, weighted reliability was 0.63, and omega was 0.52 ([Table S20](#)).

The lowest individual-level reliability omega was above one suggested cutoff (0.50) but not the higher suggested cutoff (0.60). Readers can judge this in conjunction with the weighted reliability and the more sophisticated measurement models that show high metrics for model fit at the individual level ([Table S15](#) and [S16](#)). We report omega analyses because we anticipate that some readers will want these tests, but we place more importance on the structural equation models and multi-level configural invariance tests. We also place importance on the ability of the scale to pick up on differences in behavior ([Supplementary Section 1](#)).

In creating the scale, we thought it was important to cover a wide range of the complex phenomena that make up collectivism. One of our core motivations in creating the responsibility scale was that it’s important to ask about specific relationships and concrete scenarios, rather than use the same vague, abstract wording across items (such as “the group” or “others”).

However, covering a wide range of items causes lower reliability scores. One research team put it bluntly: “The broader the construct the lower the fidelity”<sup>3</sup>. Measuring just one narrow slice of collectivism will improve reliability, but it sacrifices coverage of the whole phenomenon. In that case, the predictive power of the measure will



most likely be worse (lower external validity across behavioral markers of collectivism). In sum, analyzing correlations between items can be useful (such as with item-total correlations), but we think it is more important to judge scales based on breadth of coverage, theoretical coherence, and markers of external validity.

## 12. Survey of Researchers

We surveyed researchers who had published articles in the *Journal of Cross-Cultural Psychology* in 2017 and 2018. We chose this journal because it is a prominent journal in the field of cultural psychology and because we can be certain that the articles are about culture and not the many other sub-fields of social psychology.

To survey researchers, we sent an email to the corresponding author of every paper published in 2017 with the goal of emailing 100 researchers. Because 2017 did not have enough authors to reach our goal, we emailed authors in 2018 until we reached our goal. As an incentive, we made a \$5 contribution to the International Association of Cross-Cultural Psychology for every completed survey.

We received 40 complete responses. [Table S14](#) shows the demographics of the researchers. [Table S14](#) also shows the accuracy broken out by researchers' backgrounds. Accuracy differed less than 10% based on researchers' seniority (such as graduate students versus full professors) or their cultural background (grew up in the West versus East Asia).

In the survey, we told researchers: "We surveyed 100 people in East Asia and 100 people in the US on a scale measuring the following items. Which culture do you predict agreed more with each item?" Researchers could choose the US, East Asia, or "about equal."

Researchers then gave their guesses for 15 items. [Fig. S11](#) shows the guesses for all items. We drew items from three classic collectivism scales<sup>2,3,85</sup>. We also included items from our responsibilism scale. We focused on warm fuzzy items because this was the crucial prediction—that cultural psychologists still hold warm fuzzy intuitions.

However, we also thought it was important to test our assumption that cultural psychologists would also endorse responsibilities in close relationships. Thus, we also included the item, "We should keep our aging parents with us at home." This item reflects responsibilism, not warm fuzzies.

The guesses for the aging parents question showed strong agreement ([Fig. S11](#)). Almost all cultural psychologists (95.1%) predicted that people in East Asia would agree more. Only 4.9% predicted "about equal," and no one predicted that Americans would agree more. In the analyses, we separate out the 14 warm fuzzy items because our theory is that people will tend to guess the warm fuzzy items incorrectly and the one responsibilistic item correctly.

**Calculating Accuracy by Average Guess:** In this section, we report different methods of calculating correct guesses in the survey. Our goal is to test whether the results are similar regardless of the method used to estimate accuracy. The results show that the majority of guesses were incorrect, regardless of how we calculate accuracy.

In the main analysis, we measured accuracy of researchers' guesses by calculating the average guess. Under this coding, we coded guesses for China as 1, "about equal" as 2,



and US as 3. Thus, higher numbers mean more people guess the US will agree more than China. Then we coded the actual data as 1 if China agreed more, 2 if it was equal, or 3 if the US agreed more.

We calculated this two ways. In one way, we compared these guesses to statistical significance. Thus, if participants in China agreed significantly more than participants in US, we coded this as 1. If there was no significant difference between the US and China, we coded this as 2. If participants in the US agreed more, we coded this as 3.

In the other way, we ignored statistical significance and looked at the direction of differences. Thus, if participants in China agreed more than participants in the US (regardless of significance), we coded that as 1. If the US and China agreed equally, we coded that as 2. If participants in US agreed more, we coded that as 3.

The benefit of calculating average guesses is that it captures the full, continuous range of people's guesses. It avoids simplifying the guesses into a simple "majority wins" coding system. However, the downside is that using average guesses does not allow us to calculate a percentage correct.

But we can calculate how the average guesses correlate with the actual results across cultures. Across the 14 warm fuzzy items, the average guess was negatively correlated with the data based on statistical significance,  $r = -0.18$ , and direction,  $r = -0.44$ . We judge these by effect size rather than statistical significance because significance would not be appropriate with aggregated guesses on a small sample of items. The guesses showed a "small" deviation from the actual data based on significance and a "very large" deviation based on direction (using recent standards for effect sizes)<sup>86</sup>. The negative correlations suggest that, on average, researchers' intuitions about warm fuzzy items are in the wrong direction from the data.

**Calculating Accuracy Comparing the Most Popular Guess to Statistical Significance:** Next, we calculated accuracy by comparing the most popular guess to statistical significance. For each question, we took the guess with the highest percentage of guesses as the key guess. If the most popular guess was the US, we coded it as correct if participants in the US agreed with that item significantly more than participants in China. If the most popular guess was China, we coded it as correct if participants in China agreed significantly more than participants in the US. If the most popular guess was "about equal," we coded it as correct if there was no significant difference between the US and China. Under this coding system, out of the 14 warm fuzzy items, 14% were correct and 86% were incorrect ([Fig. S11](#)).

**Calculating Accuracy by Direction:** Finally, we can measure whether the guesses were correct by ignoring statistical significance and comparing the direction of the cultural difference. In this system, if more people guess that people in the US will agree with an item more than people in China, we code it as correct if participants in the US agree more (regardless of statistical significance). The downside of this coding system is that it ignores the "about equal" guesses. Under this system, the researchers scored slightly higher. For the 14 warm fuzzy items, guesses were correct for 21% of the items and incorrect for 79% of the items.

### 13. Testing Earlier Collectivism Scales for Divergent Validity

One basic goal of this study was to test whether the responsibilism scale is measuring something different from earlier collectivism scales. We tested this in three ways:

1. **East-West Differences:** The first way we tested this is to see whether the responsibilism scale could reliably detect East-West differences, which the earlier scales failed to do, as evidenced in several reviews<sup>12,13,11</sup>.
2. **Individual-Level Correlations:** Another way we tested this is to have participants take the responsibilism scale and other collectivism scales to see if they correlate. If responsibilism is truly different from what the previous collectivism scales were measuring, then the scales should not be strongly correlated.
3. **Culture-Level Predictive Power:** Finally, we tested this by collecting data across cultures and testing (a) whether the previous scales correlate with responsibilism at the culture level and (b) whether cultures scores using the previous collectivism scales predict external markers of collectivism as strongly as responsibilism.

**Individual-Level Correlations:** Because the full responsibilism scale is quite long, it was not feasible to give participants in 100 cultures multiple collectivism scales. Thus, we ran specific sub-samples of at least 100 participants to take the responsibilism scale and one of eight prior collectivism scales.

We decided which collectivism scales to include based on Oyserman and colleagues' meta-analysis<sup>11</sup>. First, we compared citation counts for the collectivism scales included in the meta-analysis. [Table S9](#) lists the citation counts. Then we tested every scale with over 1,000 citations (a total of six scales).

We excluded Schwartz's value scale, even though it has over 1,000 citations. We excluded this scale because it is not designed to measure collectivism narrowly. Instead, it measures values broadly, similar to the World Values Survey. We report Hofstede's scale in [Supplementary Section 27](#) rather than the main text because of critiques that it lacks face validity.

In addition to the scales with over 1,000 citations, we tested two more scales:

- We tested the more recent interdependent self scale from Vignoles and colleagues<sup>45</sup>. We chose this scale because it received extensive validation testing in 33 cultures.
- Finally, we tested Oyserman's collectivism scale because a member of the research team often uses this scale.

Our goal was to collect at least 100 participants for each collectivism scale. We collected more participants for the Singelis horizontal-vertical collectivism scale and Cross's relational self-construal scale because several collaborators use these scales in their own research.

It would have been unwieldy to give all eight scales (with over a hundred questions) to all participants. Instead, most participants who completed other collectivism scales only completed one collectivism scale besides the responsibilism scale. Thus, one

limitation is that we do not have data on how the other collectivism scales correlate with each other, only how responsibilism correlates with each scale. [Table S27A](#) lists the sample sizes and data sources for each of the prior collectivism scales.

**Culture-Level Predictive Power:** At the culture level, we tested the predictive power of responsibilism versus the five most-cited collectivism scales ([Table S9](#)). For the Singelis 1994 scale, we found a published study with scores for 39 cultures that also have responsibilism scores<sup>33</sup>. We then collected original data from the next four most-cited collectivism scales: Singelis et al., 1995, Cross et al., 2000, Gudykunst et al., 1996, and Earley, 1993.

We tested these four earlier collectivism scales by collecting data across 50 cultures. We used this data to estimate culture scores, which we compared to external markers of collectivism. Our main question was whether responsibilism predicts external markers of collectivism more strongly than previous collectivism scales.

The set of scales here is smaller than in the individual-level tests because it was not feasible to collect data across so many cultures with all eight previous scales. However, we believe that comparing responsibilism to the five most-cited collectivism scales is a comprehensive test of responsibilism versus the established scales.

As in the main analyses, we focused on the collectivism sub-scales for the three scales that had separate individualism and collectivism sub-scales (both Singelis scales and the Gudykunst scale). Testing just the collectivism sub-scales allowed us to include more scales in the cross-cultural data collection while avoiding overburdening participants with too many questions and compromising data quality. For example, including the individualism sub-scales of the Singelis et al., 1995 scale would require us to add 16 questions, which is almost as many questions as the Cross scale and Earley scale combined (19 questions).

Our goal was to collect at least 50 college students from each culture. College student samples have the downside of not being nationally representative samples. However, collecting college students had two benefits: (a) it made the samples more comparable to our responsibilism samples and (b) it made the samples more comparable across cultures.

We also wanted to collect data from diverse world regions beyond the West. Thus, we collected multiple cultures from Asia, the Middle East, Africa, and South America. Because our goal was to test the predictive power of the scales compared to external markers of collectivism, we focused on cultures that have data for these external markers (such as China) rather than cultures that lack data for these markers (such as Macau and Cyprus).

We collected samples using three methods. We recruited participants on the platform Prolific, the platform Toloka, and through collaborators in different cultures. [Tables S27B-D](#) list the sample demographics for each culture. On Prolific, we could pre-screen for current students. Toloka does not allow this type of custom pre-screening for surveys that use Qualtrics, so we had to collect a larger sample and restrict it to students afterwards. The initial sample included 6,254 participants, of which 4,646 reported being current students.

We used the same data-quality checks as with responsibilism, which are listed in [Supplementary Sections 15-17](#). We excluded people who were born outside or grow up

outside of the culture we were sampling, people whose IP address was not in the culture, and people whose cultural background was not from the local culture, such as students from the US studying abroad in Japan. These restrictions excluded 749 participants. We also excluded people who showed evidence of careless responding as measured by intra-individual response variability or a total survey duration of under the 10<sup>th</sup> percentile ( $N = 567$ ). [Supplementary Sections 15-17](#) discuss these criteria in more detail.

This left a total of 3,330 participants. Culture scores were highly correlated with or without the exclusion criteria. For example, collectivism scores on the Singelis et al., 1995 scale were correlated above 0.95 before and after exclusion criteria:  $r(48) = 0.96$ ,  $P < 0.001$ , 95% CI [0.93, 0.98]. Thus, culture scores were highly stable with or without the screening criteria.

As with responsibilism, we corrected for response style differences across cultures by subtracting average agreement across items of all four scales. [Supplementary Section 18](#) describes this in more detail. (We could not do this for the Singelis 1994 scale because we did not have individual-level data from the prior study.)

If we want to separate out response styles from actual cultural differences, we need a balance of positively coded and negatively coded items. The responsibilism scale was approximately balanced between positively and negatively coded items. In contrast, the previous collectivism scales were mostly positively-coded items. The four scales we ran had 43 positively coded items and only 6 negatively coded items. Thus, we equally weighted the positive and negative items to calculate average agreement.

The scales generally had acceptable reliability values. At the individual level, the weighted reliabilities (omegas in parentheses) were 0.78 (0.74) for the Singelis et al., 1995 scale, 0.82 (0.81) for the Gudykunst scale, 0.63 (0.57) for the Earley scale, and 0.86 (0.84) for the Cross scale. The Earley scale had the lowest reliability probably because (a) it had the fewest items and (b) the items covered a broader sample of situations and ideas than, for example, the Cross scale. Culture-level weighted reliability (omegas in parentheses) was: 0.93 (0.65) for the Singelis et al., 1995 scale, 0.90 (0.88) for the Gudykunst scale, 0.91 (0.58) for the Earley scale, and 0.88 (0.84) for the Cross scale.

#### **14. Is Responsibilism Related to Relational Mobility and Tightness?**

We tested whether responsibilism is related to relational mobility and tightness-looseness<sup>18,20</sup>. These two scales are newer than the collectivism scales, and they seem to reliably pick up on East-West differences. Relational mobility and tightness are correlated with collectivism, although researchers disagree about whether it makes more sense to think about these as *facets* of collectivism or completely different concepts<sup>18,20</sup>. Both scales are also a part of the movement away from self-report and toward “culture report.”

The tightness-looseness scale asks people about whether the social norms in their society are stricter or looser<sup>18</sup>. For example, in cultures like Japan and Pakistan, people agree that, “People in this country almost always comply with social norms” and “there are very clear expectations for how people should act in most situations.” In contrast, Brazil and New Zealand are looser, and people agree with these statements less<sup>18</sup>. We compared culture-level responsibilism scores to norm tightness scores from a previous

study<sup>54</sup>. Responsibilistic cultures had tighter norms,  $r(49) = 0.55$ ,  $P < 0.001$ , 95% CI [0.32, 0.72].

Relational mobility measures differences in freedom and choice in relationships<sup>20</sup>. In cultures with high relational mobility such as the US, people rate their society as having “many chances to get to know other people.” In contrast, people in cultures like Japan and Lebanon agree with societal descriptions like, “Even if these people were not satisfied with their current relationships, they would often have no choice but to stay with them.” We used relational mobility scores from the 2018 international study<sup>20</sup> plus imputed values for 21 neighboring cultures validated in our prior study<sup>53</sup>. Responsibilistic cultures tended to have lower relational mobility,  $r(53) = -0.53$ ,  $P < 0.001$ , 95% CI [-0.70, -0.31].

## 15. Data-Quality Check: Latitude and Longitude

We used two methods to check for data quality. First, we used data on participants’ latitude and longitude to check that they were in the culture we were testing. The Qualtrics survey platform estimates participants’ latitude and longitude using their IP addresses. We excluded participants who were not in the latitude or longitude of the culture we were testing ( $N = 451$ ).

We do not have IP addresses in all locations. In some sites, it was more feasible to take the questionnaire on paper, so we do not have IP addresses. In Cyprus, the National Bioethics Committee required us to turn off the feature to collect IP addresses, so we do not have location estimates for participants in Cyprus.

Of course, latitude and longitude data is not perfect. There can be legitimate reasons that people would have IP addresses outside of their culture:

- For example, people could be using virtual private networks (VPNs) to protect their privacy or get around firewalls. VPNs route people’s internet connection through servers in other locations, sometimes in other countries.
- Participants could also be on vacation.
- Latitude data is estimated and therefore can contain mistakes.

However, we thought excluding people who are outside the culture would improve data accuracy. To test this, we calculated responsibilism scores for each culture before and after excluding non-local IP address. Then we tested how those responsibilism scores correlated with seven external markers of cultural differences in collectivism:

- i. Global Collectivism Index<sup>15</sup>
- ii. Kinship Intensity Index<sup>16</sup>
- iii. Blood donations to non-family (lower in collectivistic cultures)<sup>16</sup>
- iv. Blood donations to family (higher in collectivistic cultures)<sup>16</sup>
- v. Parking tickets from UN diplomats in New York City<sup>32</sup>
- vi. Expectations that a friend would ask you to lie in court to avoid punishment<sup>16</sup>
- vii. An index of outgroup discrimination across societies including factors such as nepotism in business and legal codes that impose restrictions on minorities<sup>87</sup>

If removing participants who have IP addresses outside the culture makes the estimates a better reflection of the local culture, the correlations with these external markers should increase. The data showed a slight improvement in external validity. After removing these participants, the correlations of responsibility with external markers of collectivism increased slightly from  $r = 0.664$  to  $r = 0.665$ .

We also compared average responsibility scores with and without the check on latitude. In Western cultures, people with incorrect latitudes tended to score higher ( $M = 58.56$ ,  $SE = 5.28$ ) than people with correct latitudes ( $M = 36.75$ ,  $SE = 0.61$ ). In non-Western cultures, people with incorrect latitudes tended to score lower ( $M = 75.68$ ,  $SE = 1.98$ ) than people with correct latitudes ( $M = 78.07$ ,  $SE = 0.33$ ).

These results suggest that using IP addresses to estimate people's latitude was at least somewhat accurate. Removing people with incorrect latitudes led to slightly higher correlations with markers of external validity. Removing people with incorrect latitudes also increased the sensitivity of the responsibility scale at detecting cultural differences.

Despite all of that, it is worth keeping in mind that the differences were very small. Excluding this small number of participants made little difference in the overall pattern of the data. In [Supplementary Section 10](#), we report results with and without restrictions that try to separate people with IP addresses or family heritages from outside the culture. Culture scores for responsibility were nearly identical with or without these restrictions,  $r(98) = 0.99$ ,  $P < 0.001$ , 95% CI [0.985, 0.993]. Thus, the conclusions of the analysis are robust to this way of analyzing the data.

## 16. Data-Quality Check: Careless Responding

We also checked for careless responding. To measure careless responding, we followed research on the use of intra-individual response variability<sup>88</sup>. The logic is similar to finding careless responders by looking for people who answer the same number over and over again. However, researchers have suggested that it is more useful to use *variability* across questions (as measured by IRV), rather than just answering the same number. The idea is that variability can pick up on lazy responding and participants who are cheating but also smart enough to know that answering the same number over and over might lead to them getting caught.

Based on prior research, we flagged participants as “careless responders” if they met any of these three criteria:

- Their standard deviation across all responsibility items was less than 1.
- Their standard deviation across the latter half of the items was less than 1. The extra attention to the later items is based on evidence that people are more likely to respond carelessly toward the end of surveys, when they get tired.
- If they answered the last five responsibility items with the same response option.

We excluded these participants from the main analyses and the calculation of culture scores. After taking into account other exclusion criteria (except response time, discussed in the [next section](#)), the careless responder criteria flagged 737 participants (5.4% of participants).



We tested whether excluding careless responders improved data quality by calculating correlations between cultures' responsibilism scores and 7 external markers of collectivistic behavior, such as the Global Collectivism Index and the Kinship Intensity Index. We calculated culture scores for responsibilism with and without careless responders. Then we saw how these culture scores correlated with the external markers of collectivism. We calculated the *absolute* correlations because most correlations were positive (such as with the Global Collectivism Index), but one correlation was negative (blood donations to non-family).

Not surprisingly, excluding a small percentage of participants made only small differences in the average scores for cultures. However, the average correlation with external markers of collectivism increased slightly from  $r = 0.654$  to  $r = 0.658$ . The difference is small, which suggests that excluding careless responders only led to slight differences in culture scores. But the slightly higher correlation with external markers of collectivism suggests that excluding careless responders led to higher-quality data.

We also found that excluding careless responders had different effects in Western versus non-Western cultures. In Western cultures, careless responders scored an average of 58.39 ( $SE = 2.05$ ) on responsibilism compared to an average of 36.28 ( $SE = 0.63$ ) for non-careless responders. In non-Western cultures, careless responders scored an average of 73.87 ( $SE = 0.88$ ) on responsibilism, whereas non-careless responders scored an average of 78.33 ( $SE = 0.35$ ).

This pattern fits with the idea that these markers of careless responding were correctly identifying people who were not answering seriously. As a result, their scores tended toward the middle, diluting cultural differences.

## 17. Data-Quality Check: Survey Duration

In addition to response variability, we used response times for the entire survey as a marker of careless responding. In this method, we followed prior research suggesting that fast response times are markers of inattentive responding<sup>89</sup>. One study compared multiple methods of identifying careless responding and concluded that "completion time is the most effective...indicator"<sup>90</sup>.

We used the index of low variability and plotted that marker of careless responding by survey completion time. The graph showed that people who completed the entire survey in under 400 seconds had much higher rates of careless responding. For example, participants who completed the survey in 150 to 200 seconds had over a 30% probability of being marked as careless responders. But above 400 seconds, the probability dropped to under 10%.

We use survey duration as a secondary marker of careless responding beyond the response variability metric. Although the two are correlated, there is evidence that survey duration identifies careless responders above and beyond other measures of careless responding<sup>90</sup>. Thus, we restricted participants who completed the entire survey in under 400 seconds (including the consent form, survey items, and demographics). This flagged 595 participants (4.6%) who were not already flagged based on response variability or other metrics.

We tested whether exclusions based on response times improved indicators of external validity and sensitivity. If abnormally fast response times are catching careless

responding above and beyond response variability, this exclusion should increase correlations of the responsibility scores with external markers of collectivism. As in the [previous section](#), we calculated the average absolute correlation with 7 markers of collectivism across cultures.

Excluding participants with abnormally low response times improved the average correlations slightly from  $r = 0.673$  to  $r = 0.675$ . The exclusions also slightly increased the size of West vs. non-West cultural differences from  $d = 4.553$  to  $d = 4.575$ . These changes are small, but they are consistent with the idea that participants with abnormally short response times on the questionnaire were responding carelessly.

We think these data-quality checks have value, but it is worth pointing out that they make very little difference in the data. For example, cultures' average scores for responsibility are nearly identical if we use all participants' data or use all of our cultural background and data-quality screening criteria,  $r(98) = 0.99$ ,  $P < 0.001$ , 95% CI [0.985, 0.993]. Thus, the conclusions of the study remain unchanged with or without these criteria.

## 18. Accounting for Response Styles

Researchers have found differences in how people tend to use response scales across cultures<sup>59</sup>. For example, people in some collectivistic cultures (but not all) tend to agree with questions more than people in individualistic cultures<sup>58,91</sup>. If people are using scales differently across cultures, this might obscure cultural differences.

We dealt with this potential for bias in several ways. First, we designed the survey with qualities that help avoid response style bias:

**Concrete scenarios:** We used concrete scenarios to make it more likely that people across cultures understood the scenarios similarly. If people have a concrete understanding of what the questions are asking, they may be less likely to rely on a general default response style.

**Positive and negative items:** We used a mix of positive and negative items, rather than having all items where agreeing means higher responsibility. With this design, if people agree with every item, they will not get a high responsibility score just because they agree with items. It also allows us to distinguish between a pattern of responsibility responses versus simply agreeing with all the items.

**Statistically accounting for response style:** We tested different methods of statistically adjusting for response style. In the simplest method, we followed a method that previous researchers have used to adjust participants' responses on each item for acquiescence<sup>45,92</sup>. The method is simple. We calculated each participant's average agreement across all of the responsibility questions (regardless of whether agreement represented more responsibility or less responsibility). Then, for each item, we calculated an adjusted response by subtracting the participant's average agreement from that particular item<sup>91</sup>. This is the calculation:

$$\text{Question1Adjusted} = \text{Question1} - \text{AverageAgreement}$$

Not all researchers agree with this method, but previous studies have found that adjusting for acquiescence like this can improve results across cultures. For example, two

studies found that adjusting for acquiescence with this calculation more reliably reproduces the Big Five factor structure in Germany and across 18 cultures<sup>91,93</sup>.

Because some researchers disagree about using this method, we tested whether this method improved the performance of the scale. To do that, we tested whether response style differences were obscuring cultural differences in two ways:

#### **A. Adjusting for Response Style Increases Correlations with External Markers of Collectivism**

We compared correlations of cultures' raw responsibility scores and adjusted scores with seven external markers of cultural differences in collectivism:

- i. Global Collectivism Index<sup>15</sup>
- ii. Kinship Intensity Index<sup>16</sup>
- iii. Blood donations to non-family (lower in collectivistic cultures)<sup>16</sup>
- iv. Blood donations to family (higher in collectivistic cultures)<sup>16</sup>
- v. Parking tickets from UN diplomats in New York City<sup>32</sup>
- vi. Expectations that a friend would ask you to lie in court to avoid punishment<sup>16</sup>
- vii. An index of outgroup discrimination across societies including factors such as nepotism in business and legal codes that impose restrictions on minorities<sup>87</sup>

We calculated the average correlation of these seven external markers with raw responsibility culture scores and the scores adjusted for response style. We calculated the absolute correlations because one marker is negative. Adjusting for response style increased the average correlations with external markers from  $r = 0.641$  to  $0.675$ .

Note that adjusting for response style only makes small differences in culture scores. The raw culture scores correlated highly with the adjusted scores,  $r(98) = 0.99$ ,  $P < 0.001$ , 95% CI [0.984, 0.993]. However, the results suggest that adjusting for response style increases the predictive power of the measurement.

#### **B. Adjusting for Response Style Reveals Larger Cultural Differences**

Next, we tested whether adjusting for response styles led to larger cultural differences. If response style differences across cultures are obscuring cultural differences, the data should show larger cultural differences after adjusting for response style. We tested this by comparing West and non-Western culture responsibility scores with and without adjusting for response style. Cultural differences were larger after adjusting for response style,  $t(66) = 16.34$ ,  $P < 0.001$ ,  $d = 4.58$ , 95% CI [3.62, 5.52], than before adjusting  $t(66) = 15.66$ ,  $P < 0.001$ ,  $d = 4.39$ , 95% CI [3.45, 5.31].

In sum, adjusting for response style increased correlations with external markers of collectivism and revealed larger differences. These results are consistent with the idea that differences in response style are adding noise that obscures cultural differences.

We are making our calculations for raw and adjusted scores open for other researchers to use. One benefit of adjusting for response style this way is that it is simple

and easy to calculate. Any researcher can calculate this adjustment with simple formulas in a spreadsheet.

At the same time, we recognize that there are more sophisticated methods to take response style into account. For example, researchers have used structural equation models to put response style as a factor into the model<sup>45</sup>. More sophisticated methods might improve the predictive power of surveys.

By comparison, our adjustments are basic. However, simplicity has the advantage of being easy to understand and easy for future researchers to use in their own studies. Despite its simplicity, this simple method increased the ability of the surveys to tap into cultural differences.

We also tested whether validity would increase if we also took into account participants' response variability (also called "dispersion"). One method to do this is to divide each participant's response by their overall standard deviation across items<sup>92</sup>. Results showed that taking dispersion into account failed to increase correlations with external markers of collectivism. [Supplementary Section 31](#) reports the full analysis.

**Using Structural Equation Modeling to Account for Response Style:** Some researchers are critical of the simpler method of taking into account response style and instead recommend a more sophisticated method using structural equation models. We did this by adding a "common-method factor" to our models of responsibilism. We tried two methods that researchers suggest: either with response style identical for every scale item (setting the loading to 1 for each item) or by allowing response style to be different for different items (called "bifactor"). If cultural differences remain similar after accounting for response style, it suggests that response style differences are not strongly affecting the data.

The results showed that the bifactor model fit the data better. [Table S17](#) shows the results of this model. We then used this model to estimate culture scores using this method to take into account response styles. These culture scores were highly similar to culture scores using our simpler method of accounting for response style,  $r(98) = 0.94$ ,  $P < 0.001$ , 95% CI [0.91, 0.96].

Next, we tested whether using the new estimates improved external validity by calculating the average correlation of the culture scores with seven societal markers of collectivism (described in [Supplementary Section 1](#)). The average absolute correlation was slightly lower for the culture scores estimated using structural equation models ( $r = 0.635$ ) than for the simpler method ( $r = 0.675$ ). Only the simple method improved external validity over raw responsibilism scores ( $r = 0.641$ ). Thus, the simpler method to take into account response style differences demonstrated better external validity.

## 19. Estimating Culture Means

We calculated responsibilism culture scores using the responses of people who were born in the culture, grew up in the culture, and from the majority cultural group. For example, if a participant grew up in China but now lived in the United States, we did not include that person's data to calculate culture scores. As another example, the culture score for Saudi Arabia did not include participants who identified culturally as Yemeni or Egyptian. There were 2,135 participants excluded from calculations of culture scores using these metrics.

Although we think it is important to take into account migration and people's experiences in other cultures, these exclusions made little difference to the results of the study. For example, we calculated culture scores for responsibilism with all participants and then after excluding these "unmatched" participants (without any of the data-quality checks below). The two culture scores correlated very highly,  $r(98) = 0.99$ ,  $P < 0.001$ , 95% CI [0.990, 0.996].

[Fig. S14](#) shows culture scores with and without these unmatched participants. The estimated scores for the cultures change very little based on these restrictions. This suggests that analysis decisions about data screening have very little effect on the results.

## 20. Prior Collectivism Scales Mix Warm Fuzzy Intuitions and Responsibilistic Intuitions

Our basic idea is **not** that people have totally wrong intuitions about collectivism. Instead, we argue that people have a mix of incorrect warm fuzzy intuitions and correct responsibilistic intuitions. You can see this mix of intuitions in the classic collectivism scales. For example, the classic Singelis collectivism scale contains three items that reflect tight ties in close relationships:

- iv. We should keep our aging parents with us at home.
- v. Children should feel honored if their parents receive a distinguished award.
- vi. Children should be taught to place duty before pleasure.

We added these three items from the vertical collectivism scale from Singelis and colleagues<sup>3</sup> to our responsibilism scale. We chose these items to go in the responsibilism scale because they reflect tight ties in specific close relationships. The external validity tests suggest that these items do differentiate collectivistic and individualistic cultures. For example cultures' average responses on these items correlated with cultures' scores on the Global Collectivism Index<sup>15</sup>:

Aging parents:  $r(97) = 0.82$ ,  $P < 0.001$ , 95% CI [0.68, 0.88]

Feel honor:  $r(97) = 0.68$ ,  $P < 0.001$ , 95% CI [0.46, 0.79]

Duty before pleasure:  $r(97) = 0.53$ ,  $P < 0.001$ , 95% CI [0.42, 0.77]

Since the responsibilism scale contains three items from the vertical collectivism scale, it is logical that the vertical collectivism scale had the highest correlations with the responsibilism scale. In the individual-level analysis, vertical collectivism correlated with responsibilism,  $r(821) = 0.41$ ,  $P < 0.001$ , 95% CI [0.35, 0.47] ([Table S9A](#)). At the culture level, the correlation was higher,  $r(48) = 0.71$ ,  $P < 0.001$ , 95% CI [0.53, 0.82] ([Table S9B](#)).

These results suggest that there is both overlapping variance and non-overlapping variance. Overlapping variance was 16.8% at the individual level and 50.4% at the culture level. In other words, about half of the variance is distinct between the two scales. The non-overlapping variance could explain why the responsibilism scale produced

higher external validity correlations (reported below), despite similarity between the scales.

**The Case of Vertical Collectivism:** The case of the vertical collectivism scale offers insight into the challenges around theorizing about collectivism. On the one hand, we think the creators of the vertical collectivism scale created some of the best collectivism items of all of the previous scales. Items like, “we should keep our aging parents with us at home” fit neatly with our criteria for responsibilism by mentioning concrete responsibilities in specific close relationships. It is for this very reason that we chose items from the vertical collectivism scale as a part of the responsibilism scale.

At the same time, the authors of the vertical collectivism scale seemed to have missed what we think is the central insight of their scale. Rather than seeing the scale as about duties and responsibilities in close relationships, the authors described it as about “status” and “inequality”<sup>3</sup>. That is why they labeled it “vertical.” They saw the scale as measuring “the way people perceive, accept, and manage inequality”<sup>3</sup>. This is despite the fact that none of the scale items use terms like “status,” “hierarchy,” or “inequality.”

This (mis)understanding of the vertical collectivism scale was not a one-off occurrence. Many later researchers have used the scale to measure attitudes toward inequality. For example, a 2018 paper used the scale and described it as measuring “the valuing of hierarchy”<sup>94</sup>. In 2021, researchers used the scale and defined it as “preferences for equality and inequality”<sup>95</sup>.

The case of the vertical collectivism scale illustrates how devilish these misunderstandings of collectivism seem to be. In a sense, the solution (in our perspective) to the collectivism puzzle was hiding in plain sight. Even when researchers had items that tap into what we think is the more accurate conceptualization of collectivism—responsibilities in specific close relationships—they built a different theory around those items. The theory and even the name “vertical collectivism” misspecified the construct.

The description of the construct seemed to miss what was truly distinct about these scale items. The answer has been in front of our eyes for years, yet the insight remains elusive. The fact that later researchers have continued to describe the scale as primarily about inequality suggests that people have very different understandings from what we are proposing about what the vertical collectivism scale is actually measuring.

**External Validity of Responsibilism Versus Vertical Collectivism:** Despite the virtues of the vertical collectivism scale, it fell short of the responsibilism scale in our tests of external validity. The three responsibilistic items in the vertical collectivism scale can help explain why it performed well on tests of external validity. These items might also explain why the horizontal collectivism scale performed so poorly. Our explanation is that the horizontal collectivism scale performed poorly because it contains items that tap into warm fuzzy intuitions, such as “To me, pleasure is spending time with others.”

Using the same method as in Fig. 6, we broke out the external validity performance of the Singelis 1995 scale out into horizontal and vertical collectivism. We calculated the correlation across seven markers of external validity ([Supplementary Section 15](#) lists the seven markers). As in Fig. 6, responsibilism had an average correlation of  $r = 0.71$ , whereas horizontal collectivism had a correlation of  $r = -0.19$ , and vertical collectivism has a correlation of  $r = 0.55$ .



We also calculated responsibilism's external validity performance without the three Singelis scale items. This allows us to test whether the high external validity performance of the responsibilism scale is primarily due to the Singelis items. Without the Singelis items, the average correlation was  $r = 0.67$ . This is similar to the correlation with the Singelis items. The similar results suggest that the performance of the responsibilism scale is not primarily driven by the Singelis items.

One argument could be that the responsibilism scale has higher external validity performance than the vertical collectivism scale because the responsibilism scale has more items (14 items versus 8 items). Having more items could allow for more precise measurements and higher external validity. We tested this by limiting the responsibilism scale to the first 8 items in the scale and re-calculating the external validity correlations. With the same number of items in both scales, responsibilism had an average external validity correlation of  $r = 0.72$ , compared to  $r = 0.55$  with vertical collectivism. This result suggests that the higher performance of the responsibilism scale is not a result of simply having more items.

We also compared the responsibilism scale and vertical collectivism scale by putting them in a regression predicting scores on the Global Collectivism Index. We chose this index because it has the data for more cultures than any other of our external validity metrics. In that regression, responsibilism predicted collectivism scores at  $\beta = 0.66$  compared to  $\beta = 0.27$  for vertical collectivism.

Another way to test the two against each other is with partial correlations. We did this by running separate partial correlations testing each scale while controlling for the other scale. We ran partial correlations for each scale with each of the seven markers of external validity. On average, responsibilism remains a "strong" (by Cohen's standards) external validity correlation after accounting for vertical collectivism ( $r_{\text{partial}} = 0.50$ ), whereas vertical collectivism drops to a "small" correlation ( $r_{\text{partial}} = 0.16$ ).

In sum, the vertical collectivism scale performs well compared to previous collectivism scales in our analysis. This performance fits with our choice to include items from the vertical collectivism scale in our responsibilism scale based on the fact that those items describe responsibilities in specific close relationships. However, the responsibilism scale shows higher external validity correlations, even after removing the vertical collectivism items, after limiting the responsibilism scale to the same number of items, and after controlling for vertical collectivism.

Furthermore, there is a meaningful amount of non-overlap between the two scales. At the individual level, the shared variance was 16.8% ([Table S9A](#)). At the culture level, the shared variance was 50.4% ([Table S9B](#)). For readers interested in the theory and measurement of vertical collectivism, we have a separate paper that discusses these questions in depth<sup>96</sup>.

## 21. Rating Self vs. Others

We wanted to know if measuring responsibilism with self-report is different from "culture report." Culture report is when researchers ask participants to rate their society, not themselves. It is important to understand whether we can still use self-report because some researchers responded to the collectivism crisis by abandoning self-report in favor

of culture report<sup>18–20</sup>. For example, one study found that self-report ratings failed to find the expected East-West differences but culture report did<sup>12</sup>.

Abandoning self-report comes with downsides. For one, culture report risks tapping into people's unfounded stereotypes. It also risks measuring pluralistic ignorance, where everyone believes *other people* hold a value, even though they themselves reject it<sup>21</sup>.

If we abandon self-report, we might also lose the opportunity to measure individual differences, such as the differences in social class. For example, the relational mobility scale uses culture report, and it found no differences based on age or socioeconomic status, despite surveying over 16,000 people<sup>20</sup>.

We tested this by randomly assigning a sub-set of participants in the US ( $N = 147$ ) and Hong Kong ( $N = 516$ ) to take the responsibility scale for themselves and for their society. The participants took both versions, but we randomly assigned some people to take the self version first and others to take the culture version first. For the culture-report condition, we used the instructions from the relational mobility scale<sup>20</sup>, which asks people to rate their society:

“Please rate how accurately these statements describe the people in the immediate society (your school, workplace, town, neighborhood, etc.) in which you live. Please indicate how true you feel each statement to be for the people around you.”

[Fig. S5](#) shows the results. Replicating the earlier difference, the self-ratings of responsibility were higher in Hong Kong than the US,  $t(661) = 12.48$ ,  $P < 0.001$ ,  $d = 1.17$ , 95% CI [0.97, 1.36]. The “culture report” ratings also showed Hong Kong was more responsibility than the US,  $t(661) = 15.47$ ,  $P < 0.001$ ,  $d = 1.45$ , 95% CI [1.25, 1.65]. Americans rated themselves as responsibility as their society, paired-samples  $t$  test:  $t(146) = -0.04$ ,  $P = 0.970$ ,  $d = 0.00$ , 95% CI [-0.17, 0.16]. Hong Kongers rated themselves less responsibility than their society,  $t(510) = -5.79$ ,  $P < 0.001$ ,  $d = -0.26$ , 95% CI [-0.34, -0.17].

In sum, the responsibility scale picked up on cultural differences in self-ratings and culture ratings. Cultural differences were larger with culture ratings, which fits with previous findings<sup>12</sup>. It is still an open question whether self-ratings or culture ratings are a better predictor of cultural differences in behavior. But the results show that both methods are useful to measure cultural differences with the responsibility scale.

If we can keep using self-report, it has the benefit of allowing researchers to test for individual differences. Researchers often want to test collectivism differences between individuals, and self-report is a more appropriate tool for that than culture report. However, we wanted to test culture report because culture report may still have advantages. For example, culture report may reveal larger culture differences if people's perceptions of social norms differ more strongly than people's perceptions of their own attitudes.

Culture report can also be useful because people don't always behave based on their own attitudes. Sometimes social norms are a stronger influence on people's behavior than their own attitudes or preferences. For example, I might want to speak directly with other people, but if I think the people around me value indirect communication, I might choose to speak indirectly. Thus, we think both self-report and culture report have value. But the

results from this study suggest that we do not need to abandon self-report to measure cultural differences.

## 22. Did the Pandemic Affect the Results?

We wanted to test whether the COVID-19 pandemic influenced the results because we collected data in some cultures before the pandemic and in other cultures after the pandemic had started.

### A. China Before and During the Pandemic

The strongest way to test whether the pandemic affected the results is to test the same people before the pandemic and during the pandemic. We did this by testing college students at Hubei University in 2019 and 2020. We were able to test 104 students across both time periods. The average age of the longitudinal sample was 18.9 years at Time 2, with 78.8% female respondents. At Time 1, the students were first-year students, and many of them had just left home for the first time. At Time 2, it was September, 2020, and they were just starting their second year at university. At that time, school was in person for these students.

Longitudinal data is a strong method, but we cannot tell apart differences from the COVID-19 pandemic and differences as people get older. College students are at an age and a stage in life where they experience lots of change. Thus, it is worth keeping in mind that this data covers a period of development change in addition to the change from before the pandemic to during the pandemic.

[Fig. S12](#) shows that responsibilism went down from 2019 to 2020, paired-samples  $t$  test,  $t(103) = 2.83$ ,  $P = 0.006$ ,  $d = 0.28$ , 95% CI [0.08, 0.47]. The size of the change is similar to changes in personality in longitudinal studies of college students. For example, a meta-analysis found the average change in conscientiousness for 18-22-year-olds was  $d = 0.22^{97}$ .

Although responsibilism went down, this change was far smaller than East-West differences. For example, in the main study, responsibilism scores were 41 in the US and 94 across sites in China. Although responsibilism went down at the Hubei site in China, it went from 96 to 88. At 88, the difference between the US and China is still large.

This result suggests that any effect of the pandemic is minor compared to East-West cultural differences. This conclusion is similar to the conclusion of a study that tested family values before and during the pandemic<sup>98</sup>. That study found statistically significant but small changes during the pandemic. The changes were small compared to the larger patterns of family values.

It is worth noting that our sample in China was in Wuhan. Wuhan was the epicenter of China's COVID-19 outbreak. Wuhan experienced more disruption, news coverage, and deaths than any other part of China. Thus, Wuhan is probably an extreme test case of change during the COVID-19 pandemic.

## B. India Before and During the Pandemic

We also collected longitudinal data in India on Amazon's MTurk platform in October, 2014 and late April, 2021. We did not plan on collecting longitudinal data, but we had the opportunity to test before and during the pandemic because the MTurk platform records worker ID numbers. With these limitations, we were able to get longitudinal data for only 28 people. This sample is very small, but we thought it was still worthwhile to present because it offered the rare opportunity to test for longitudinal change over an extraordinarily long period of time. We did not include this data in the main analysis because MTurk workers are primarily adults past college age, which is different from the samples we collected in other cultures.

Time 2 was during an intense period of the pandemic in India. Late April was near the peak of India's crippling delta wave, which overwhelmed hospitals and [killed 240,000 people](#).

[Fig. S13](#) shows responsibility scores in India from 2014 to 2021. Responsibility did not differ significantly over time, paired-samples  $t$  test,  $t(27) = 1.54$ ,  $P = 0.136$ ,  $d = 0.29$ , 95% CI [-0.67, 0.09].

### 23. Longitudinal Test-Retest Reliability

The longitudinal data can also speak to the question of longitudinal test-retest reliability. However, it is worth keeping in mind that the COVID-19 pandemic is an extreme circumstance to test longitudinal reliability in. The disruption of the pandemic may increase the amount of change over time.

Scores from China in 2019 correlated with scores in 2020,  $r(102) = 0.40$ ,  $P < 0.001$ , 95% CI [0.22, 0.55]. In India, scores from 2014 had a similar correlation with scores in 2021, although the correlation was not significant, perhaps because the sample was small,  $r(26) = 0.36$ ,  $P = 0.058$ , 95% CI [-0.01, 0.65]. This correlation is similar to longitudinal correlations of Big Five personality scores over time. A meta-analysis found that test-retest correlations averaged around  $\rho = 0.50$  among 18-22-year-olds<sup>99</sup>.

It is worth noting that these tests are longitudinal over time. This is different from test-retest reliability measured on the same day or the same week. Correlations would probably be higher if we were to test people more closely in time.

### 24. Definition of West, Sometimes West, and Non-West

For international comparisons, we split cultures into Western, "sometimes Western," and non-Western. Many cultures are easy to categorize: the US and Germany are Western. China and India are non-Western.

But there are many cultures that are harder to categorize. Is Russia Western? Are Latin American cultures like Colombia and Brazil Western? Some definitions of West emphasize the legacy of Christianity, which would include Colombia and Brazil as Western. Yet other definitions recognize the cultural differences between Latin American cultures and Western Europe.

We started with the broad outlines of Samuel Huntington's definition ([map archive](#)). This definition categorizes countries as Western, non-Western, or a third category in between. We call that third category "sometimes Western." The "sometimes Western"

category includes Latin America, the Caribbean, and Eastern Europe. We then made three edits to Huntington's map:

- i. To define Eastern Europe, we used the extent of the Soviet Union bloc ([map archive](#)). This stretched from Russia as far west as Poland, Czechoslovakia, Hungary, and Yugoslavia. This categorization is similar to earlier studies that have contrasted former Soviet Union bloc areas with non-Soviet bloc areas in Western Europe<sup>18,45</sup>. Even within Germany, there are still cultural differences between the former democratic Western areas and areas that used to be a part of the Soviet Union, such as looser social norms in the Western area<sup>18</sup>.
- ii. We counted southern European cultures as "sometimes Western" based on studies finding that southern European cultures like Greece, Italy, and Spain are more collectivistic than northwest European cultures like Germany and the UK<sup>16,31</sup>. Our categorization corresponds to the United Nations geoscheme definition of southern Europe ([map archive](#)). Several countries in this scheme were also in the Soviet bloc (such as Bosnia). But the UN geoscheme adds Portugal, Spain, Italy, and Greece.
- iii. Huntington classified Greenland as Western, probably because it is a territory of Denmark. We classified Greenland as non-Western because of its majority Inuit cultural heritage and language.

Differences in responsibilism suggested that there is a valid distinction between cultures in western Europe versus southern Europe and the former Soviet Bloc. The 18 former Soviet Bloc and southern European cultures in our sample scored higher on responsibilism ( $M = 49.21$ ,  $SE = 2.23$ ) than the 13 Western European cultures ( $M = 26.83$ ,  $SE = 3.52$ ),  $t(29) = 5.64$ ,  $P < 0.001$ ,  $d = 2.05$ , 95% CI [1.15, 2.93].

The data showed that this simple three-category structure explained a large portion of the variation in responsibilism around the world. We can compare how useful these three categories are versus using all 100 separate cultures. To do that, we ran one-way ANOVAs predicting individuals' responsibilism scores from the three categories or by splitting all 100 cultures out separately. Then we compared how much variation in responsibilism the three categories versus all 100 cultures could explain.

Splitting out all 100 cultures explained 32.2% of variation in responsibilism. Splitting cultures into just these three groups explained 24.9% of variation. In other words, placing cultures in these three groupings capture more than half of the variation that using all 100 cultures captured. In sum, categorizing cultures together will always miss some real variation between cultures, but these three categories capture much of the variation between cultural groups. This suggests that the three categories of West, "sometimes West," and non-West are meaningful categories, explaining a large portion of cultural variation.

It is important to remember that the main conclusions of our study do not depend on this categorization. Scholars have argued about how to define Western culture. Analysis results show that this operationalization is empirically useful, but it is open for debate. The central tests of this study are (a) that responsibilism reliably detects cultural differences that previous reviews failed to find and (b) these differences correspond to

real-world behaviors that reflect collectivism. West vs. non-West categorizations are not necessary for either of these empirical tests.

## 25. Cultures Versus Countries

We use cultures as our unit of analysis rather than countries because cultures do not always align with country borders. Here are two examples. Puerto Rico is a US territory, but it is majority Spanish-speaking with nearly 400 years as a colony of Spain. Greenland is a territory of Denmark, but it is majority Inuit and speaks Greenlandic.

We treated the Puerto Rico sample as a distinct culture from the mainland US, and we treated the Greenland sample as a distinct culture from Denmark. The data suggests it is accurate to treat these as distinct cultures. Puerto Rico scored more responsibilistic (45.7) than the mainland US (40.9). Greenland scored more responsibilistic (55.8) than Denmark (0.0).

Of course, there will always be debate about what should count as separate cultures. In the same way, there will always be debate about what should count as dialects versus separate languages. However, in identifying cultures, we looked for three factors:

i. Different majority languages (for example, Puerto Rico is majority Spanish, whereas the US is majority English; Greenland speaks Greenlandic, but Denmark speaks Danish).

ii. Different governments, passports, political systems (for example, although Hong Kong is a territory of China, its passports are different from the Mainland and allow for more visa-free international travel than Mainland passports)

iii. Differences in historical government (for example, Macau is a territory of China, but it was under Portuguese rule for over 400 years)

There are surely cultural differences beyond the 100 cultures we present scores for. We describe a few in this study, such as between the US south and north. However, we encourage other researchers to use our open data to explore other cultural differences that we do not cover here.

## 26. Scale Creation: Avoiding English-Centric Bias

In developing the scale, we wanted to avoid having it be based solely on English. Our perception is that psychologists develop scales in English first, set the scale, and then translate it directly into other languages. That means that many questionnaires in psychology are based on English primarily, without sensitivity for how natural the items sound in other languages. For example, a scholar of Chinese talked about how some phrases in Chinese have an air of “translatese”<sup>100</sup>. The idea is that some phrases sound like English just written in Chinese.

We wanted to avoid this problem. To do that, we avoided first setting the wordings in English without considering other languages. Instead, we developed the scale in both English and Chinese at the same time. We followed these steps:

- i. The first author wrote the scale items in English.



- ii. The first author (who is fluent in Chinese) translated the items into Chinese and checked the translations with a native speaker.
- iii. When wordings sounded unnatural, stiff, formal, or simply hard to translate into Chinese, we changed the *English* wordings until they were easy to translate into Chinese in a way that sounded natural.

In many cases, this process made the questions less abstract, less formal, more concrete, and free from idioms. These goals fit with our original goal of making the scale less abstract and more concrete, which a previous study found improved the ability of questionnaires to measure cultural differences<sup>10</sup>.

Chinese has the benefit of being a very different language from English—a different language family and writing system. This should make the scale more free of English-centric bias.

We think developing a scale in multiple languages at once is an improvement over how psychologists usually create scales, but our process was not perfect. For example, English and Chinese are only two languages. Both English and Chinese mostly lack linguistic features like gender and case. Future versions can improve on the scale by editing the wordings to be more equivalent with languages that have gender and case, such as Hindi, German, and Russian.

## 27. What About Hofstede's Cultural Dimensions?

Hofstede was an early pioneer of using self-report surveys to measure cultural differences<sup>46</sup>. His surveys of IBM employees around the world in the 1960s established cultural dimensions that researchers still use today, including in our research<sup>101</sup>. Hofstede's early data found differences between East and West, which suggests it might have more external validity than many of the classic collectivism scales.

However, we report tests of Hofstede's individualism survey here rather than in the main text because we believe understanding this scale benefits from more context. Our basic argument is that items in the Hofstede scale lack face validity. One example is the item that asked employees to rate the importance of "good *physical working conditions* (good ventilation and lighting, adequate work space, etc.)."

How would an item about lighting and ventilation make it into a collectivism scale? To create the scale, Hofstede used bottom-up analyses<sup>46</sup>. In a sense, his analyses looked for survey items that correlated with each other, and then he created cultural dimensions from those correlations.

This bottom-up approach has strengths. Letting the data guide the analysis can alert us to the incorrect assumptions we start out with. A more bottom-up approach might have prevented the problems with warm fuzzy items in collectivism scales.

However, the bottom-up approach can also make mistakes. Just because two survey questions correlate with each other does not guarantee they are measuring the same idea. To avoid this problem, we have to balance the bottom-up approach with theory and careful reflection.

We believe that Hofstede's bottom-up approach produced a jumble of items that did not all tap into individualism. This was probably compounded by the fact that Hofstede's original surveys were designed for the purposes of IBM's offices, not cultural psychology. The theoretical jumble is clear in the survey item about lighting and ventilation, which had the third highest factor loading of all the items.

We have tried asking our colleagues to guess whether agreeing with the importance of ventilation makes people more individualistic or less individualistic. People find it hard to guess. In Hofstede's data, agreeing about lighting and ventilation makes people less individualistic. This item was not a one-off accident. It remained in some future Hofstede survey waves ([web archive](#)).

We suspect caring about lighting and ventilation was correlated with individualism because of wealth. Wealthy Western countries tend to be individualistic<sup>15</sup>. And IBM employees in wealthy countries like Switzerland and Canada probably didn't think much about lighting and ventilation because those are not serious problems there. But employees in developing countries like Bangladesh or Ghana probably thought about lighting and ventilation more because they could not take these things for granted.

Poverty could also explain other puzzling items from Hofstede's scale. For example, the highest-loading item was about wanting a job that "leaves you sufficient time for your personal or family life." Agreeing with this makes people individualistic in Hofstede's scale. But poverty offers a plausible explanation. If people are struggling to get by, free time is less important than getting a higher-paying job.

The high correlation between poverty and collectivism can probably explain why Hofstede's individualism scores are so widely used—they produce the correlations researchers expect. We think the problem is that Hofstede's scores produce those correlations for the wrong reasons.

Understanding this background is helpful when assessing external validity correlations. External validity correlations cannot tell the difference between (1) scales that work because they're effectively tapping into the cultural construct and (2) scales that work because they're tapping into a highly correlated third variable. External validity will tell us that both of these types of scales are performing well.

This logic applies to the classic [correlation between ice cream sales and shark attacks](#). Ice cream sales might help us predict shark attacks, but we know that ice cream sales are not measuring shark attacks.

With that context in mind, here we report external validity correlations of Hofstede's scores. We calculated the average correlation with the seven markers of external validity (such as blood donations to strangers, listed in [Supplementary Section 15](#)). For this analysis, we limited the sample to cultures that have data for both responsibility and Hofstede's scores<sup>46</sup>. To make the numbers go in the same direction, we flipped the correlations for Hofstede's individualism scores from negative to positive.

On average, Hofstede's individualism scores predicted the external markers in the theoretically consistent direction. The average correlation for Hofstede's individualism was  $r = 0.62$ , compared to  $r = 0.67$  for responsibility.

## 28. Social Class Analysis

In the main text, we report that responsibilism correlates with social class, whereas other collectivism scales do not. We report the correlations in Fig. 5 and for all of the sub-scales in [Table S12](#).

If we assume that the “true” correlation is the correlation that responsibilism has with social class, the effect size is small:  $r = 0.11$ . That means the samples with some of the other collectivism scales are too small to detect that correlation with certainty. If we assume a correlation of  $r = 0.11$ , a sample would need 646 participants to achieve 80% statistical power.

Although the samples for the other collectivism scales are too small to detect a correlation of 0.11, the results across the scales are still helpful. For one, the correlations across the different scales give an overall picture of the correlations. Averaging over the different scales increases the sample size, yet the correlation is still in the opposite direction as the result for responsibilism. With the responsibilism scale, people of high social status are less collectivistic. With the other collectivism scales, people of high social status are slightly more collectivistic.

Second, the “true” correlation could be larger for the other scales, and the sample sizes could then be big enough to detect those effects. Thus, we think there is value in presenting the results for all the scales, but we recognize the limited sample sizes.

## 29. Percentage of the World Population Covered

In the main text, we report the percentage of the world population covered by the cultures we sampled. We calculated this based on population data from the World Bank (<https://data.worldbank.org/indicator/SP.POP.TOTL>). Because most data came from 2019, 2020, and 2021, we calculated the percentage statistics for each year. The results are slightly different depending on which year of data we use for the population statistics, but all are over 90%.

Year 2019: 90.20%

Year 2020: 90.20%

Year 2021: 90.11%

## 30. Accounting for Response Variability Did Not Improve Predictive Power

In [Supplementary Section 18](#), we reported that adjusting for participants’ average agreement across items improved the predictive power of the responsibilism scores. Here, we tested whether accounting for differences in participants’ response *variability* could further improve predictive power. This could help correct for extreme response style. The idea is that participants who use more extreme responses (and therefore have a higher standard deviation) will have their scores reduced.

To do this, we followed the method of dividing by each participant’s standard deviation across items<sup>92</sup>. First, we calculated the standard deviation across all responsibilism items for each participant. Then we divided the response to each question by each participant’s standard deviation. We did this on top of subtracting the average agreement (described in [Supplementary Section 18](#)). The calculation is:

$$\text{Question1Adjusted} = (\text{Question1} - \text{AverageAgreement}) / \text{Standard Deviation}$$

Next, we calculated culture scores for responsibilism (a) with just the average agreement adjustment and (b) adding the standard deviation adjustment. Then we calculated correlations of these responsibilism scores with external markers of collectivism.

Adjusting for standard deviation slightly decreased correlations with external markers of collectivism from  $r = 0.661$  to  $0.656$ . Before adjusting for dispersion, differences between Western cultures and non-Western cultures were  $t(66) = 16.34$ ,  $P < 0.001$ ,  $d = 4.58$ , 95% CI [3.62, 5.52]. After adjusting for dispersion, differences between Western and non-Western cultures were  $t(66) = 14.89$ ,  $P < 0.001$ ,  $d = 4.17$ , 95% CI [3.27, 5.06]. The difference between West and non-West responsibilism scores was not significant before versus after adjusting for dispersion, paired samples  $t$  test,  $t(99) = -0.28$ ,  $P = 0.780$ ,  $d = -0.03$ , 95% CI [-0.22, 0.17].

In sum, adjusting for response variability did not improve the predictive power of the responsibilism survey. The data was not consistent with the idea that response variability was obscuring cultural differences. Thus, we did not add the adjustment for response variability in the analyses in this study.

### 31. Spatial Auto-Correlation: Controlling for Latitude and Longitude

In this study, we found that responsibilism predicts differences in external markers of collectivism, such as family size and returning wallets to strangers. We also found correlations with plausible ecological causes of responsibilism, such as rice farming. However, one limitation of testing correlations between environmental variables is that they might be correlated not because of cause and effect but because the two variables tend to happen near each other in physical space (called “spatial auto-correlation”).

For example, if we were to measure shark bites and ice cream sales across the United States, we might find a correlation between the two. But we can be pretty sure that ice cream sales do not cause shark bites. The two are correlated only because they tend to be clustered together in space. People buy ice cream at the beach, and shark bites happen at the beach.

One way to check whether spatial auto-correlation is falsely inflating correlations is to run analyses that take into account latitude and longitude. We did this by re-running the correlations in Extended Data Tables 1, 2, and 4 using partial correlations that control for latitude and longitude. [Tables S30](#) and [S31](#) report the results.

Overall, results were similar after controlling for latitude and longitude. Most variables that were significant with simple correlations remained significant after controlling for latitude and longitude. For example, historical exposure to the Western Church was still significantly correlated with responsibilism,  $r(90) = -0.71$ ,  $P < 0.001$ , 95% CI [-0.81, -0.50]. Historical herding also continued to predict responsibilism  $r(96) = -0.27$ ,  $P = 0.009$ , 95% CI [-0.45, -0.04]. The fact that these correlations remained significant after taking into account latitude and longitude suggests that these correlations are not just false relationships produced by being near each other in physical space.

However, it is important to remember that controlling for latitude and longitude does not produce “true” correlations. Controlling for spatial auto-correlation can produce truer, smaller correlations, or it can produce smaller correlations for things that are

clustered together in space for real reasons. For example, diseases tend to be more common around the equator than farther north<sup>102</sup>. Controlling for latitude would probably reduce the correlation between disease and heat, which is a real cause of disease.

Thus, it is important to know that controlling for latitude and longitude can rule out false correlations based on spatial clustering alone, but it does not produce “true” correlations. Researchers still need to use judgment to draw conclusions about correlations before and after controlling for spatial clustering.

### 32. Scale Refinement Using a Holdout Sample

We tested the reliability of our scale refinement process by randomly selecting half of the participants in the dataset; running the same scale refinement process (outlined in [Supplementary Section 6](#)); and then testing the results on the other half of the sample. If the results differ meaningfully, it suggests the original results may be overfitted. If the results are similar to the earlier results, it suggests that the solution is stable. The data file randomly selected sub-sample, analysis syntax, and results output are all available on the [OSF page](#).

**Item selection:** First we ran the same scale refinement process compared the items that survived the refinement process to the items that survived the original process with the whole scale (Table S4). The analysis with the random sub-sample led to the same 14 items as the original refinement. This result suggests that the item-selection result is stable.

**Reliability:** Next, we tested the reliability of the scale on the holdout sample. The reliability was similar (weighted reliability culture level = 0.89, individual = 0.64) to the original analysis with the full sample (culture = 0.89, individual = 0.63).

**External validity:** Following that, we tested the external validity performance of the culture scores based on the holdout sample. We tested external validity by calculating the average correlation of responsibilism culture scores with seven markers of collectivism across cultures (outlined in [Supplementary Section 15](#)). In the original full sample, the average correlation was  $r = 0.68$ . In the holdout sample, the average correlation was  $r = 0.67$ . Note that these correlations are slightly different from Fig. 6 because Fig. 6 is limited to the cultures that have data from the classic collectivism scales. The analysis here uses all cultures.

**Conclusion:** In sum, the analysis with the holdout sample suggests that overfitting was not a serious problem in the original analysis. The results suggest that the scale refinement process produced reliable results.

### 33. Research Articles Mentioning “Collectivism” Increased 190% in the 1990s

In the introduction, we wrote that there was a boom in research on collectivism after the classic paper from Markus and Kitayama came out in 1991. Here, we report a small test to try to quantify this idea. Our goal was to get a rough measure of psychology research on collectivism over the decades. To do that, we ran Google Scholar searches for “psychology collectivism” and “psychology.” Then we calculated the percentage of “psychology collectivism” results per “psychology” results. We did this for each decade from 1950 up to 2019.

The results showed a spike in the 1990s around the time of the Markus and Kitayama paper ([Fig. S17](#)). Results for “collectivism” roughly doubled from the 1980s to the 1990s. Results jumped another 170% in the 2000s.

Of course, this analysis is limited. For one, it is limited to a few keywords. Google Scholar is also not a perfect record. But it is broadly consistent with the idea that there was a boom in research on collectivism.

### **34. Robustness to Even Sample Sizes and Samples of 40 Participants per Culture**

One limitation of the samples is that they are uneven between cultures. Many cultures have between 50 and 100 participants, but some cultures have several hundred participants (Table S6). In this section, we re-run several of the main analyses of the paper using equal samples across cultures. By equalizing sample sizes across cultures, we were also able to answer another question: Are the results robust to sampling just 40 or 50 participants per culture?

#### **i. Randomly sampling within cultures: Estimating the correlation with the Global Collectivism Index**

To equalize the samples, we took the entire sample (before any data-quality checks) and randomly sampled 40 participants within each culture. We chose 40 because the cultures with the smallest samples (such as the Bahamas) have enough participants for random samples of 40 participants. With that random sample, we calculated each culture’s average responsibilism score.

Next, we tested the correlation between the newly created responsibilism scores and the Global Collectivism Index, which is the external validity marker with data for the most cultures (Extended Data Table 1). We repeated this process 50 times to see how robust the correlation estimates would be.

[Fig. S18A](#) shows the distribution of the results. The average correlation across 50 iterations was  $r = 0.74$ . This is similar to the estimate in the main text of  $r = 0.75$ . Perhaps more important is the spread of the estimates. Across the 50 iterations, the standard deviation was 0.01. The most extreme estimates were 0.71 on the low end and 0.77 on the high end. Thus, the spread held quite close to the estimate in the full dataset. This result suggests that samples of 40 participants per culture are sufficient (even though not ideal) to estimate responsibilism cultures scores in ways that demonstrate a correlation with the Global Collectivism Index as a marker of external validity.

#### **ii. Randomly sampling within cultures: Estimating external validity correlations**

We then repeated that process for all of the external validity markers in the original analysis (Fig. 5, Extended Data Tables 1-2). We plotted the average correlation in the 50 random sub-samples with the original correlation with the full dataset ([Fig. S18B](#)). The gray dots are the average correlations in the random sub-samples. The peach-colored dots are the correlations with the full sample.

The closeness of the gray dots and peach dots show that the results from the random sub-samples are similar to the correlations in the full sample (Fig. 5). In some cases, the dots are overlapping so much that it is difficult to tell the two dots apart. The four widest distributions are the markers of external validity that have the smallest number of cultures (such as the economic games, which have data for just 15 cultures). This analysis



suggests that results are stable across random samples and robust to forcing the sample sizes to be equal for all cultures.

### iii. Randomly sampling within cultures: Estimating responsibility culture scores

Next, we wanted to ask how robust the estimates of responsibility are to different sample sizes across cultures. To do this, we ran random samples ranging from 5 participants per culture to 100 participants per culture. As an extreme example, we randomly sampled 5 participants within each culture and estimated responsibility scores for each culture. Then we repeated this process 50 times. After that, we did the same thing with random samples of 20 participants per culture, 50 participants, and 100 participants. To sample 100 participants per culture, we had to limit the sample to exclude cultures with fewer than 100 participants.

[Fig. S19](#) shows violin plots of the results. The gray lines stretch out to the most extreme estimates for each culture. The black lines show  $\pm 1$  standard deviation. If the estimates are normally distributed, the black lines should include approximately 68% of the estimates.

The results show a few patterns:

**5 Participants per Culture:** First, sampling 5 people per culture produces a large range of results (top-left graph on [Fig. S19](#)). That's not a surprise because 5 people is a very small sample. But it provides a baseline for comparison. With just 5 people per culture, the average standard deviation across samples is 14.4. That's about 14% of the range for the culture scores, which span from approximately 0 to 100 ([Table S22](#)).

**20 Participants per Culture:** Second, perhaps more surprisingly, the estimates narrow quite a bit with samples of 20 people per culture (top-right graph on [Fig. S19](#)). The standard deviation shrinks to 6.5.

**50 Participants per Culture:** With 50 participants, the standard deviation is 3.1. In other words, with samples of 50, 68% of estimates should fall within  $\pm 3.1$  points of the average. And 95% of estimates should fall within  $\pm 6.2$  points of the average.

With 20 or 50 participants, the graphs show that it's fairly easy for any particular culture to move up or down a few ranks. Mexico could jump above Venezuela, or Barbados could fall below the Bahamas. But it would be essentially impossible for Australia and Egypt to trade places. The relative stability of the estimates here fits with the results for external validity correlations in [Fig. S18B](#).

**100 Participants per Culture:** If we move up to samples of 100 people, the standard deviation is 1.6 (bottom-right graph on [Fig. S19](#)). That means 68% of estimates should fall within  $\pm 1.6$  points, and 95% of estimates should fall within  $\pm 3.2$  points.

We can also look at the samples that have the most variation in their estimates. The largest standard deviations were in the UK and Canada. That is probably because these are some of the most culturally diverse samples, and we ran these simulations before any data restrictions. That means the random samples in the UK and Canada could pick up on people with non-local cultural backgrounds, such as international students. If we put the sample restrictions in place, the standard deviation should be even smaller. Yet even without sample restrictions, 95% of estimates should fall within  $\pm 3.2$  points.

**Conclusion:** With samples of 50 participants per culture and no data screening, 95% of responsibility estimates fall within  $\pm 6.2$  points of the average. With samples of 100 participants, that range falls to  $\pm 3.2$  points. Thus, it is easy for cultures to jump up or down a few spots in the rankings, but the overall pattern of culture estimates seems relatively robust to samples of 50 participants.

### 35. What Counts as a Close Relationship?

Our brief definition of responsibility is responsibilities, duties, and roles in close relationships, but weaker ties in distant relationships. This raises the question of what counts as a close relationship. In this section, we discuss what counts as a close relationship.

Our survey items lay out some basic concrete examples of close versus distant relationships ([Table S4](#)). For example, several items mention parents, children, “several friends,” and a “good friend.” Items about more distant relationships mention strangers, a classmate you don’t know from a class of 100 people, and a coworker you’ve never met before. The fact that these items produce results in line with our hypotheses suggest that people tend to interpret these examples of relationships as close versus distant. These results provide some answers but also leave some open questions:

**i. Large group membership is not sufficient:** The items about unknown coworkers and the unfamiliar classmate suggest that large group membership is not sufficient to count as a close relationship in responsibility cultures. At the very least, simply knowing that two people belong to the same large group is not enough to predict that they will interpret this as a close relationship.

**ii. Do groups depend on context?** We say “not sufficient” because we think there may be situations where large group membership might count as a close relationship if there are other contextual factors. For example, a stranger on the streets of Beijing probably counts as a distant relationship. But in the context of a war, that stranger might count as a close relationship. Or if two strangers encounter each other in a public protest that they feel passionately about, that large group membership might be sufficient. Our data cannot answer these questions. Future research can explore these questions by asking a wider range of questions about different relationships and contexts.

**iii. Does family always count?** Our items assume that family relationships count as close relationships. However, there could be exceptions. For example, if someone is estranged from their father, they might not consider this relationship close. Similarly, cultures may differ in whether they think cousins, great aunts, and great uncles are close.

**iv. Letting participants define the relationships:** Some items essentially let participants define what a close relationship is by stating “good friend.” Thus, participants can imagine a good friend of their choice. However, many of our items try to provide concrete information about the relationship so that participants imagine the same thing across cultures. For example, the words “acquaintance” and “colleague” might bring to mind different things for different people. Some items in our scale give more concrete descriptions, such as “You are working in a company for two years, and one day you meet a co-worker from a different department that you’ve never met before.” Note that we previously reported this item in an earlier paper<sup>28</sup>.

v. **Avoiding assumptions of warm feelings:** One thing we wanted to avoid is assuming that relationships have to be based on warm feelings or trust. There is evidence that people can feel tied to relationships where they feel suspicion<sup>64,103</sup>. Researchers from Western cultures tend to live in social environments with higher relational mobility, where people feel reporting having more choice and freedom in their relationships<sup>20</sup>. In a high-mobility environment, it would make sense if people think good feelings and trust are pre-requisites for close relationships. But this assumption may be less common in responsibilistic cultures, which tend to have lower relational mobility ([Fig. S3](#)).

Thus, we think it is wise to avoid defining close relationships as based on warm feelings, choice, or trust. Instead, we speculate that concrete characteristics may be more useful to describe close relationships, like time spent together, shared tasks, and interdependence. However, the data in our study cannot answer all of the questions about what counts as a close relationship. Our propositions here can provide hypotheses for future studies that ask about responsibilities across different types of relationships in different contexts.

### 36. Evidence That People in Individualistic Cultures Tend to Agree More with Warm Fuzzy Collectivism Items

Our argument is that warm fuzzy collectivism is conceptually different from how people think about individualism, but it is paradoxically more common in individualistic cultures. Why is it conceptually different? Researchers often describe individualism as a separation from other people<sup>2,47</sup>. Many dictionary definitions also emphasize separation from other people, such as the *Oxford English* [definition](#): “the habit or principle of being independent and self-reliant.” Thus, we think it would be confusing if we say that a concept that emphasizes warm feelings and pro-sociality (warm fuzzy collectivism) is individualism.

However, our argument is that warm fuzzy collectivism is actually more common in individualistic cultures. How does that make sense? This could be because the lesser emphasis on responsibilities and commitment frees people up to enjoy relationships. Duties serve a purpose, but they’re not primarily about having fun. In this section, we review empirical evidence for the claim that people in Western, individualistic cultures tend to agree more than people in non-Western, collectivistic cultures with warm fuzzy collectivism scale items.

**Factor Structure:** The most central piece of evidence is in the factor structure of the responsibilism scale itself. We intentionally designed some items to describe pro-sociality and warm intentions toward distant relationships. For example, items describe trusting strangers, sharing information with an unknown co-worker, and casually starting a conversation with an unfamiliar classmate ([Table S4](#)). If warm fuzzy collectivism and responsibilism are really just the same thing (as people’s intuitive guesses suggest, [Fig. S11](#)), then cultures that agree with responsibilistic items should also agree with warm fuzzy items. In contrast, if warm fuzzy and responsibilism are opposites, the factor structure should show that the items about warm fuzzies go in opposite directions from the responsibilistic items. Results from tests of the factor structure find that the warm fuzzy items load negatively, whereas responsibilistic items load positively ([Fig. S16](#)).

**Criterion Validity Correlations for Responsibilism Scale Items:** Another critical piece of empirical evidence for this claim is in tests of how each item in the scale

correlates with external markers of collectivism, such as living with extended family. The right side of [Table S4](#) shows the external validity correlations for each item. The results are consistent with the factor structure results: items that describe pro-sociality with acquaintances or distant relationships are **negatively** correlated with external markers of collectivism. For example, the item about caring for aging parents has an average correlation with external markers of collectivism of  $r = 0.62$ , whereas the item asking about sharing information with an unfamiliar co-worker has a negative correlation,  $r = -0.55$  (note that [Table S4](#) uses reverse-coding for the warm fuzzy items, so all correlations are positive).

**Criterion Validity and Cultural Differences with Items from the Classic Collectivism Scales:** Another way to test this idea is to look at particularly warm and fuzzy items from the classic collectivism scales. Do these items correlate negatively with external markers of collectivism? For example, take the item “To me, pleasure is spending time with others.” Culture scores on this item correlate negatively with scores on the Global Collectivism Index:  $r(48) = -0.42$ ,  $P = 0.002$ , 95% CI [-0.63, -0.16]. We can also ask whether people in non-Western cultures tend to disagree with these items. To test that, we used our simple categorical variable of cultures as Western (coded as 0), “sometimes Western” (coded as 1), and non-Western (coded as 2; [Supplementary Section 24](#) explains this categorization in more detail). Non-Western cultures tended to disagree with the “pleasure” item:  $r(48) = -0.51$ ,  $P < 0.001$ , 95% CI [-0.69, -0.27]. In other words, people in non-Western (and presumably less individualistic) cultures are disagreeing with this warm fuzzy item.

Results are similar for the item “My relationships with others are more important than my accomplishments”<sup>104</sup>. Culture scores on this item correlate negatively with the Global Collectivism Index,  $r(48) = -0.67$ ,  $P < 0.001$ , 95% CI [-0.80, -0.48]. Scores also correlate negatively with our non-Western categorization variable:  $r(48) = -0.65$ ,  $P < 0.001$ , 95% CI [-0.79, -0.46].

Not all warm fuzzy items show negative correlations. For example, [Table S13A](#) shows responses to several warm fuzzy items in China and the US. People in the US agree significantly more with some items (8 out of 14 warm fuzzy items in [Table S13A](#) and [Table S13B](#)), but some items show no significant difference (4 items), and a small portion (2 items) show significantly lower agreement in the US. In other words, the overall balance of the evidence is that people in Western cultures tend to endorse warm fuzzy items more than people in non-Western cultures (with some exceptions), and warm fuzzy items tend to correlate negatively with external markers of collectivism.

### 37. The Pipeline Project Two Replication

In the main text, we briefly described the Pipeline Project Two replication of our study. Here, we describe it in more detail. The Pipeline Project Two team contacted the first author asking if he was working on any projects that would be a fit for the Pipeline Project. The goal of the Pipeline Project is to run replication studies, but before the study is actually published. This builds on the earlier Pipeline Project<sup>105</sup>. The authors argued that pipeline replications are valuable because the method “increases transparency and ensures that what appears in our journals has already been independently validated. [Pipeline replication] avoids the cluttering of journal archives by original articles and replication reports that are separated in time, and even publication outlets, without any

formal links to one another. It also avoids the adversarial interactions that sometimes occur between original authors and replicators”<sup>105</sup>.

The main hypothesis we pre-registered was that China or other typically collectivistic cultures would score higher than the US on responsibilism. The Pipeline Project collected data at six sites in the US ( $N = 554$ ), one in Australia ( $N = 260$ ), one in Scotland ( $N = 276$ ), and one in Singapore ( $N = 306$ ). One limitation is that most of the data came from Western cultures.

The results replicated the finding that typically collectivistic cultures (in this case, Singapore) score higher than the US. We pre-registered the US as the comparison, although Singapore scored higher than all three Western cultures.

The replication authors also ran an additional analysis to test whether each culture scored above or below the midpoint of the scale. This was not pre-registered or approved by us. We have never advocated for interpreting cultural differences on the scale as being about differing from the midpoint of the scale. We do not consider this a meaningful test of our hypothesis. For example, if one culture scores a 4.1 on the scale, and another culture scores a 6.1, we would interpret that as a meaningful cultural difference, regardless of whether they are significantly different from the midpoint.

However, Singapore did score significantly above the midpoint. The pooled Western sites also scored significantly below the midpoint. Looking at individual Western sites, Scotland and two US sites scored significantly below the midpoint. The other sites were not significantly different from the midpoint when analyzed individually. Thus, the pooled pattern of results using this “midpoint test” is consistent with the idea that Western cultures score lower on responsibilism than non-Western cultures, although we do not advocate for comparisons to the midpoint.

### 38. Checking Assumptions of Statistical Tests

We ran analyses in R (version 4.3.2), MPlus (version 1.8.8), and SPSS (versions 29-31). In this section, we report tests of the assumptions of the statistical tests. We checked for skewness, kurtosis, and equality of variance. First, we tested assumptions for correlation tests. We checked for skewness and kurtosis and compared these to cutoffs of  $< 2$  for skewness and  $< 7$  for kurtosis<sup>106</sup>.

Tables S33 and S34 report the skewness and kurtosis statistics for culture-level variables. Table S35 reports the results for individual-level variables. All variables were below the cutoffs, except for death rate from drug use disorders, Eastern Church exposure, and population density in 1CE. After log transformation, all of these variables were below the cutoffs.

**Death Rate from Drug Use Disorders Log:** skewness = 0.54, kurtosis = -0.30

**Eastern Church Exposure Log:** skewness = 0.88, kurtosis = -0.82

**Population Density in 1CE Log:** skewness = -0.39, kurtosis = -0.30

Next, we re-ran the correlations between responsibilism and these three indicators after log transformation. Two of the three correlations were significant in the original data and remained significant after. One correlation was non-significant and remained non-significant after transformation.

### Death Rate from Drug Use Disorders:

- **Log:**  $r(96) = -0.57, P < 0.001, 95\% \text{ CI } [-0.69, -0.42]$
- **Raw data:**  $r(96) = -0.41, P < 0.001, 95\% \text{ CI } [-0.56, -0.23]$

### Eastern Church Exposure:

- **Log:**  $r(92) = -0.50, P < 0.001, 95\% \text{ CI } [-0.64, -0.34]$
- **Raw data:**  $r(92) = -0.21, P = 0.048, 95\% \text{ CI } [-0.39, -0.002]$

### Population Density in 1CE:

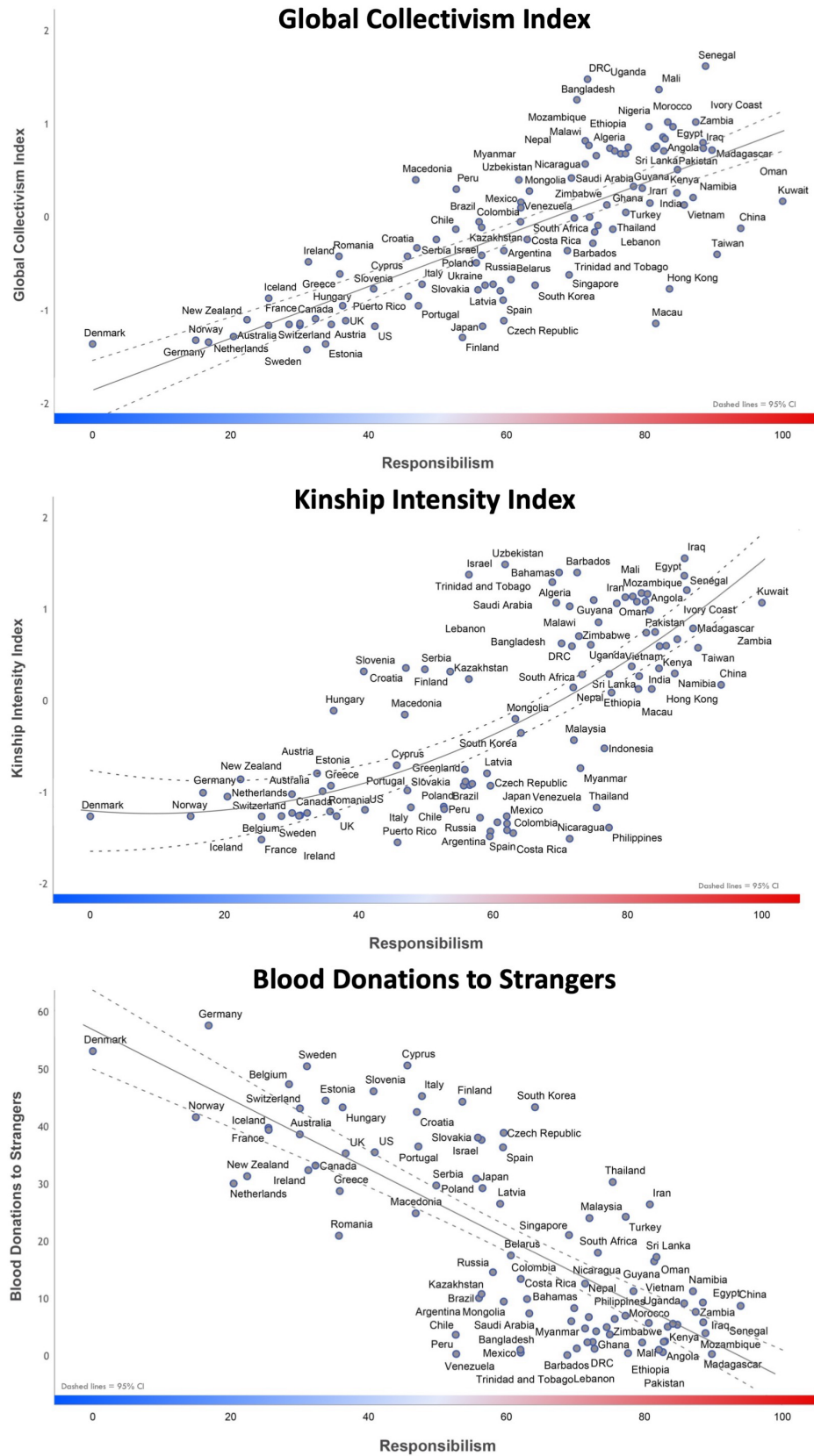
- **Log:**  $r(80) = -0.09, P = 0.446, 95\% \text{ CI } [-0.30, 0.13]$
- **Raw data:**  $r(80) = -0.21, P = 0.057, 95\% \text{ CI } [-0.41, 0.01]$

Next, we checked the assumptions of the  $t$  tests. We tested for equality of variance for each independent-samples  $t$  test using Levene's test for equality of variances. The test checks whether the variance is unequal between groups. If variance is unequal, the recommendation is to adjust the degrees of freedom in the  $t$  test.

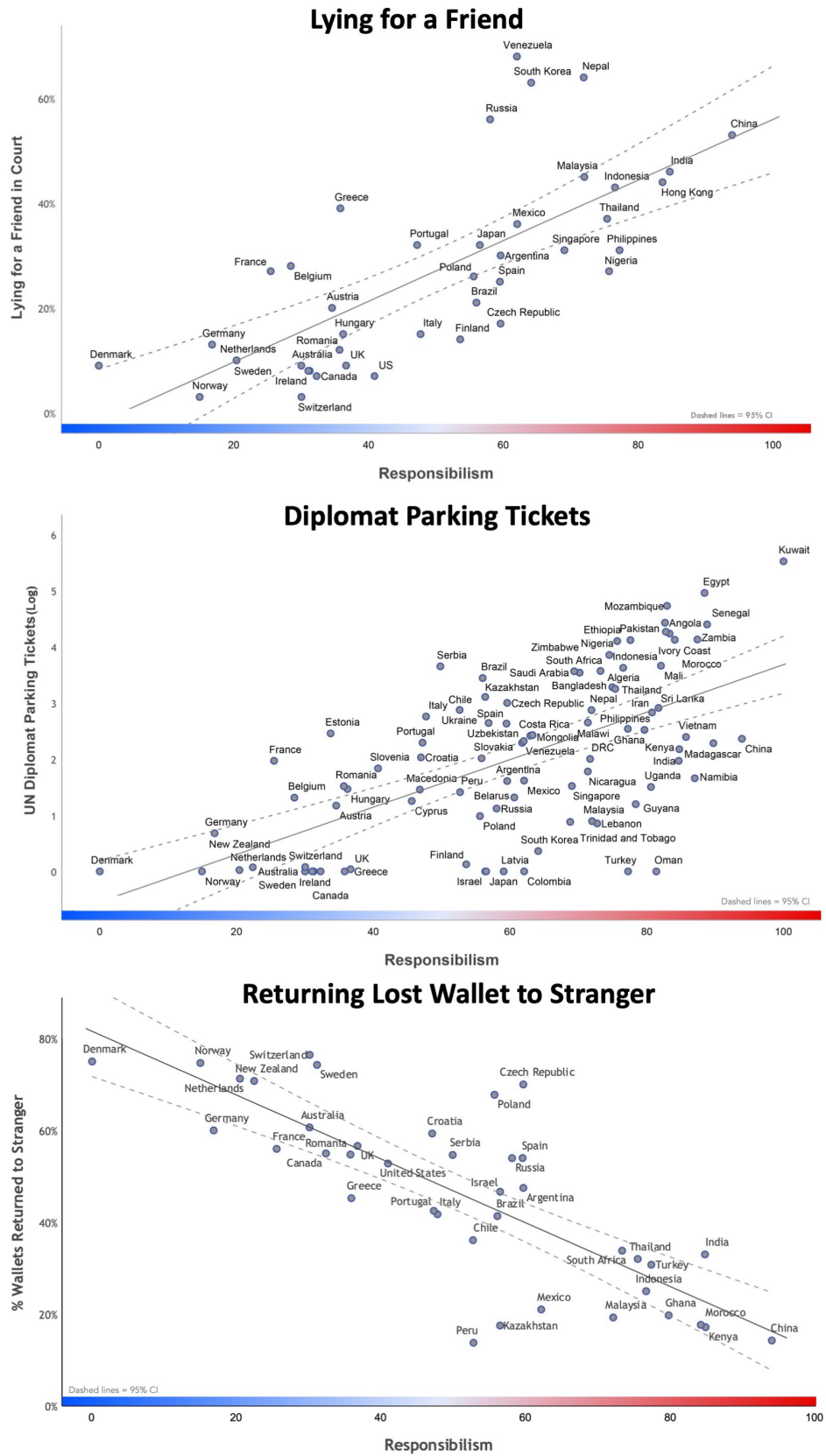
Table S36 reports the results for the tests of equality of variance. The tests turned up 8 comparisons with unequal variance. For these comparisons, we ran the analyses again after adjusting the degrees of freedom. All of the results remained significant.

- **Testing Cultural Differences without Translation:** US vs. Hong Kong:  $t(710.65) = 19.80, P < 0.001, d = 1.38, 95\% \text{ CI } [1.23, 1.54]$  (Main Text: Methods Section)
- **Self-Report Responsibilism:** Hong Kong vs. US:  $t(195.04) = 10.65, P < 0.001, d = 1.17, 95\% \text{ CI } [0.97, 1.36]$  ([Supplemental Section 21](#))
- **“Culture-Report” Responsibilism:** Hong Kong vs. US:  $t(201.02) = 13.60, P < 0.001, d = 1.45, 95\% \text{ CI } [1.25, 1.65]$  ([Supplemental Section 21](#))
- **The well-being of my coworkers is important to me:** US vs. China:  $t(112.16) = 8.10, P < 0.001, d = 1.32, 95\% \text{ CI } [1.00, 1.63]$  ([Table S13A](#))
- **To me, pleasure is spending time with others:** US vs. China:  $t(148.01) = 2.93, P = 0.004, d = 0.44, 95\% \text{ CI } [0.15, 0.74]$  ([Table S13A](#))
- **I have subdued my personal opinions in order to be part of my group:** US vs. China:  $t(184.37) = 2.14, P = 0.034, d = 0.29, 95\% \text{ CI } [0.01, 0.58]$  ([Table S13B](#))
- **One should live one's life independently of others:** US vs. China:  $t(189.97) = -4.37, P < 0.001, d = -0.60, 95\% \text{ CI } [-0.89, -0.30]$  ([Table S13B](#))
- **I would feel comfortable telling an unfamiliar co-worker about a sales account:** US vs. China:  $t(175.20) = 2.61, P = 0.010, d = 0.37, 95\% \text{ CI } [0.08, 0.66]$  ([Table S13B](#))

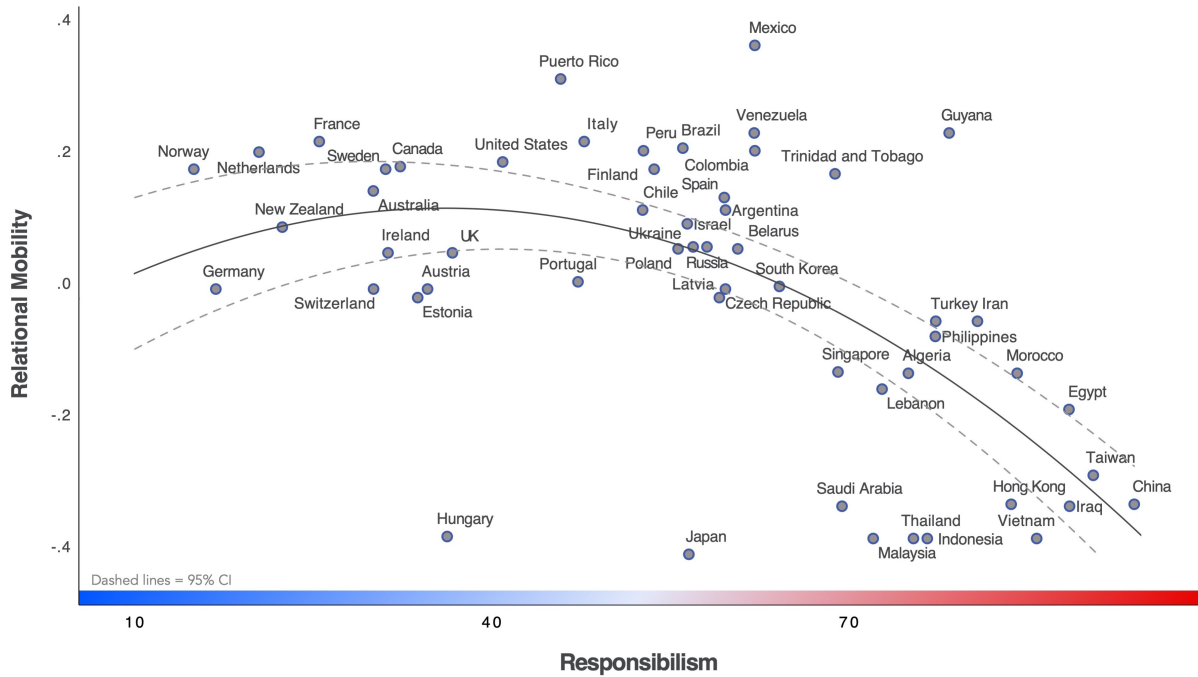




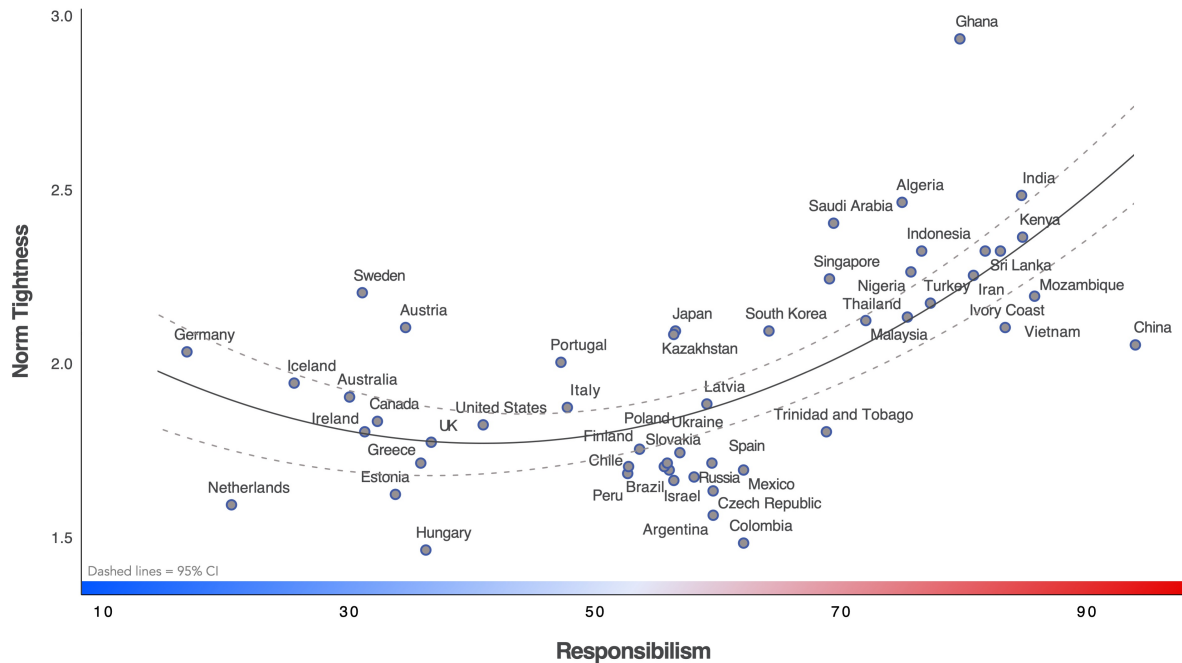
**Fig. S1. Correlations between responsibilism and collectivism markers.** Extended Data Tables 1-2 explain each measure and source.  $N = 99$  cultures, except blood donations  $N = 89$ .



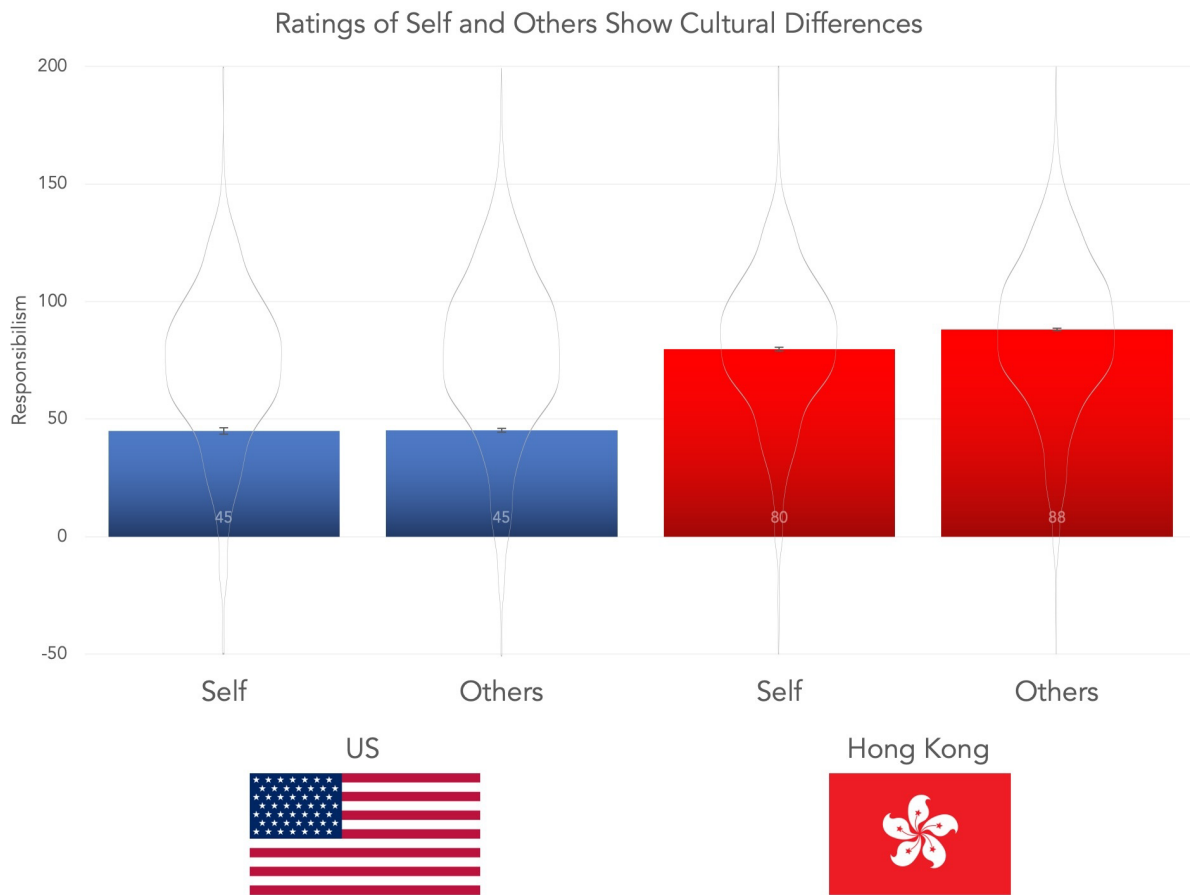
**Fig. S2. Correlations between responsibilism and collectivism markers.** Extended Data Tables 1-2 explain each measure and source.  $N = 40$  (top), 89 (middle), 39 (bottom) cultures.



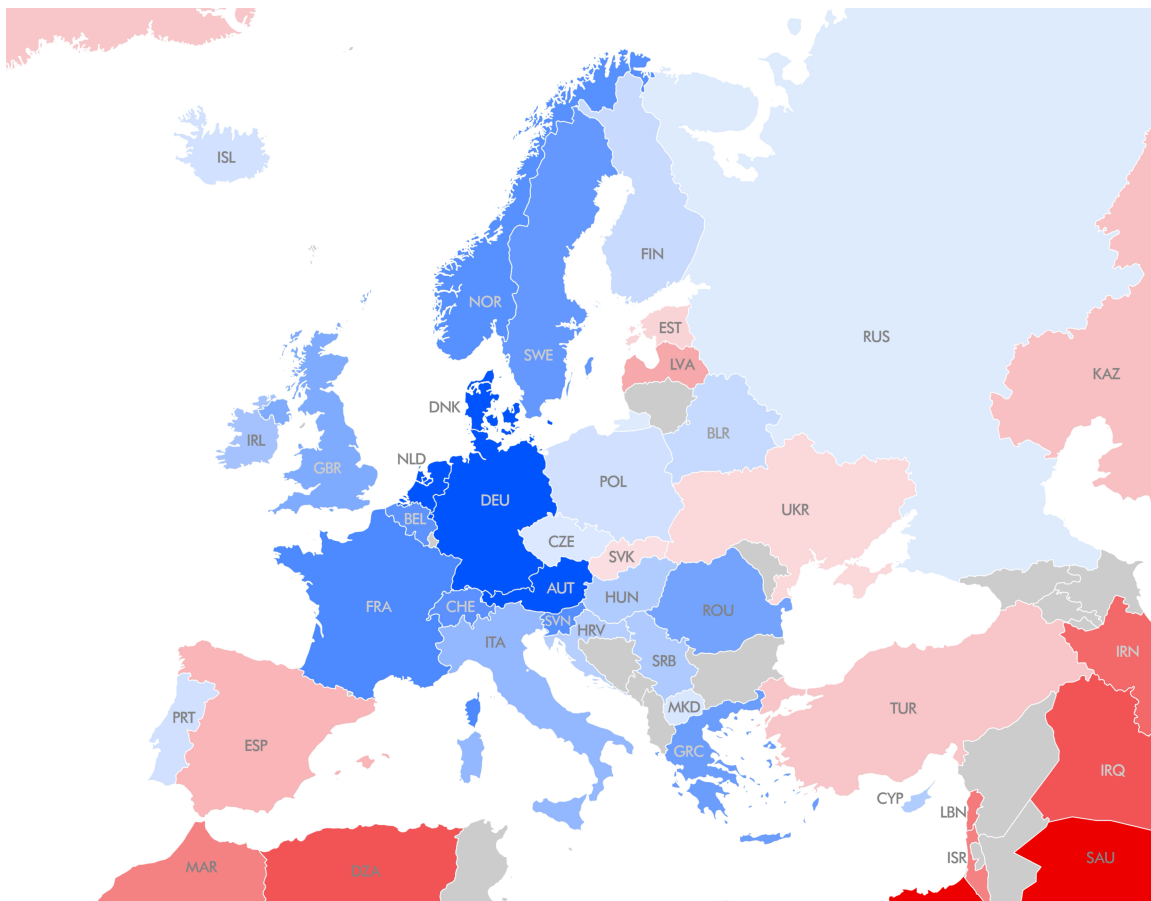
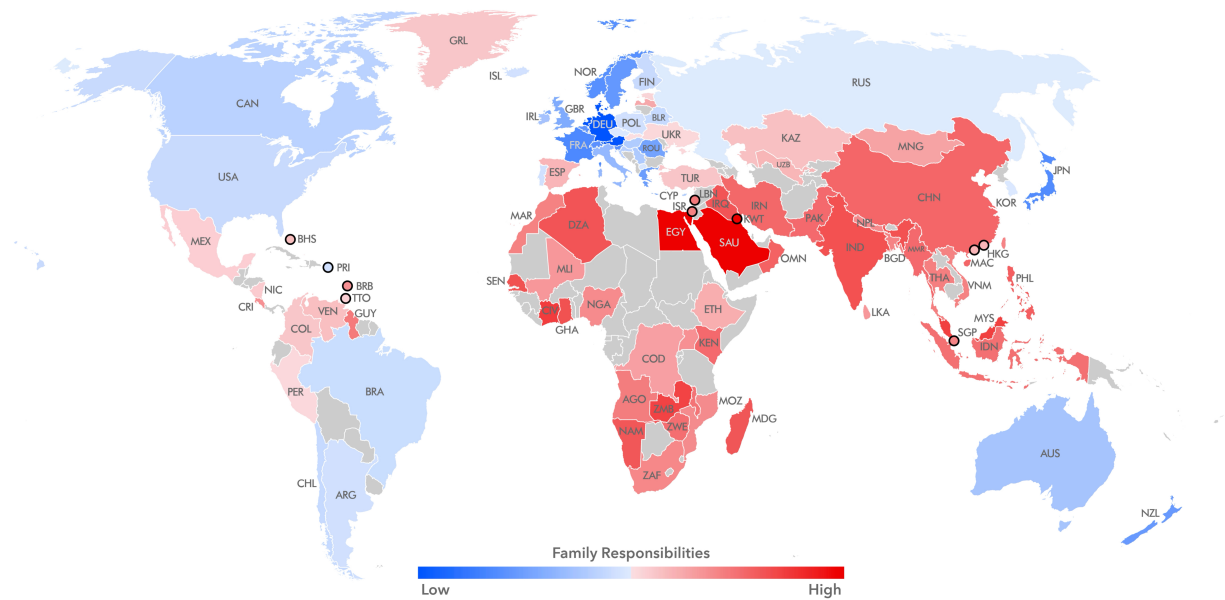
**Fig. S3. Responsibilistic cultures have lower relational mobility.** Cultures with strong responsibilities in close relationships tend to have less freedom and turnover in their relationships  $r(53) = -0.53$ ,  $P < 0.001$ , 95% CI [-0.70, -0.31]. The graph uses a quadratic fit line because the fit improved from an R squared of 0.28 for a linear fit to 0.37 for a quadratic fit. Relational mobility data comes from a 2018 study<sup>20</sup> with additional neighboring cultures imputed and validated from a later study<sup>53</sup>.  $N = 55$  cultures.



**Fig. S4. Responsibilistic cultures have tighter social norms.** Cultures that emphasize duties in close relationships tend to have more binding social norms,  $r(49) = 0.55$ ,  $P < 0.001$ , 95% CI [0.32, 0.72]. The graph uses a quadratic fit line because the fit improved from an  $R$  squared of 0.30 for a linear fit to 0.44 for a quadratic fit. Tightness values come from a 2021 study<sup>54</sup>.  $N = 51$  cultures.

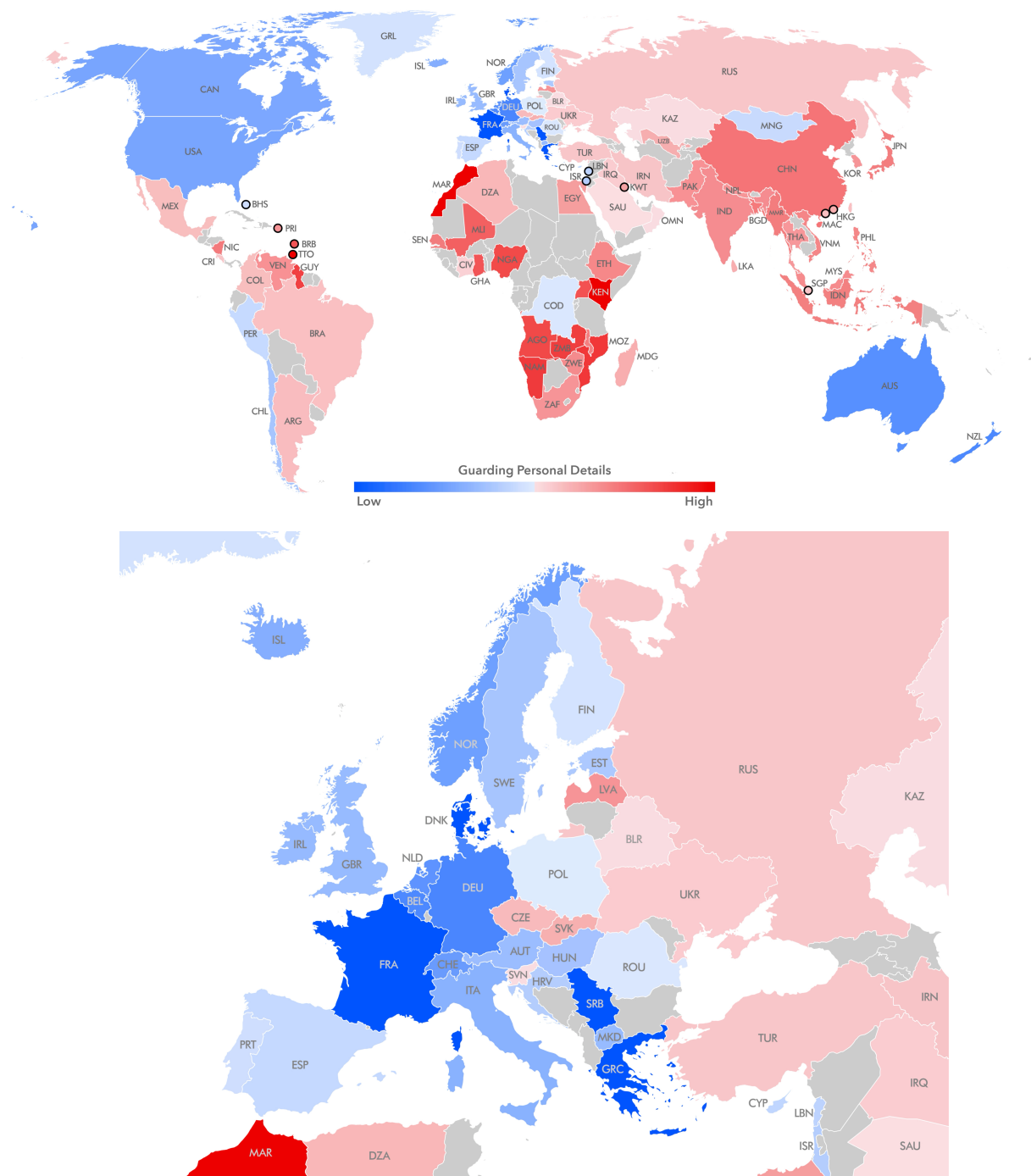


**Fig. S5. Responsibility Ratings for Self-Report vs. “Culture Report” in the US and Hong Kong.** We randomly assigned participants in the US and Hong Kong to rate themselves on the responsibility scale (as normal) or rate their society. Participants took both versions, but we randomized some participants to take the self version first and others to take the society version first. In the “culture report” scale, we asked participants to, “Please rate how accurately these statements describe the people in the immediate society (your school, workplace, town, neighborhood, etc.) in which you live. Please indicate how true you feel each statement to be for the people around you.” Both rating methods showed cultural differences, but cultural differences were a little larger with culture ratings. Error bars = 1 SEM. Scores can be below 0 because 0 corresponds to the lowest average score for a culture (Denmark). Of course, an individual can score below the average for Denmark.  $N = 663$  participants.

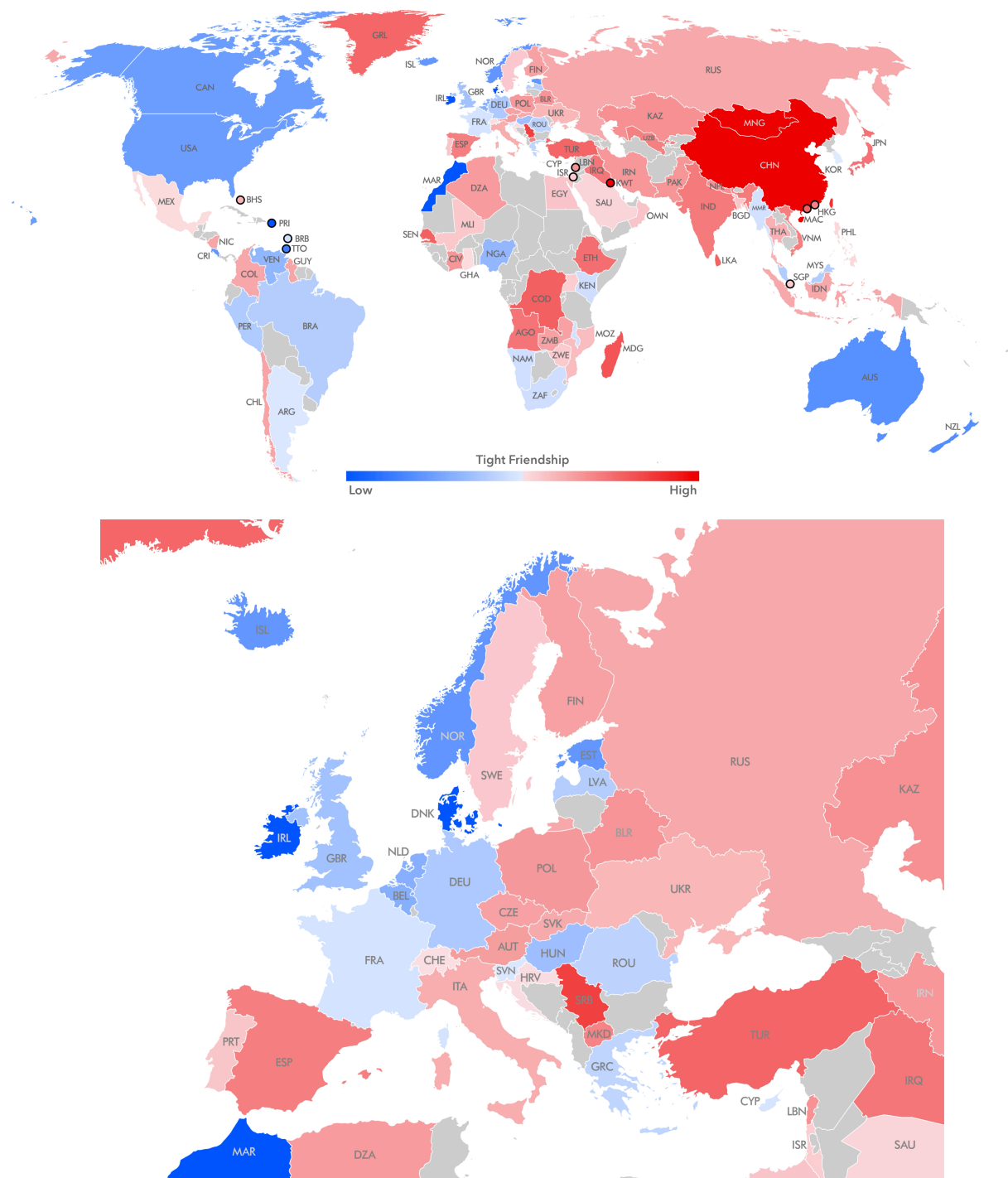


**Fig. S6. Responsibilism Component Scores for Family Responsibilities.** Cultures in red scored higher on family responsibilities. Responsibilistic cultures tended to score higher on family responsibilities. Items on this component have to do with duties and taking care of family members, such as “We should keep our aging parents with us at home.” This component is calculated from questions 1, 6, and 8 ([Table S4](#)).

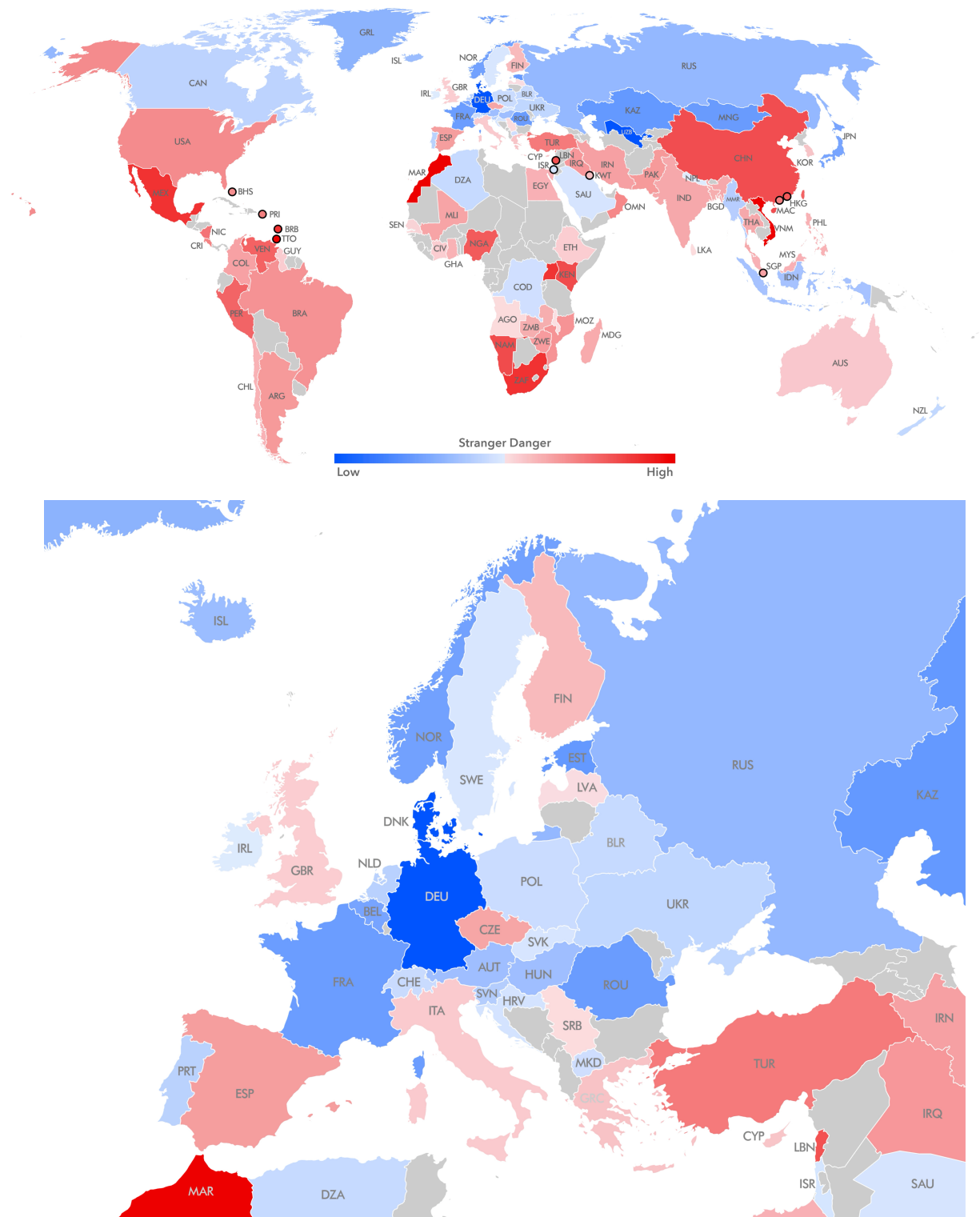




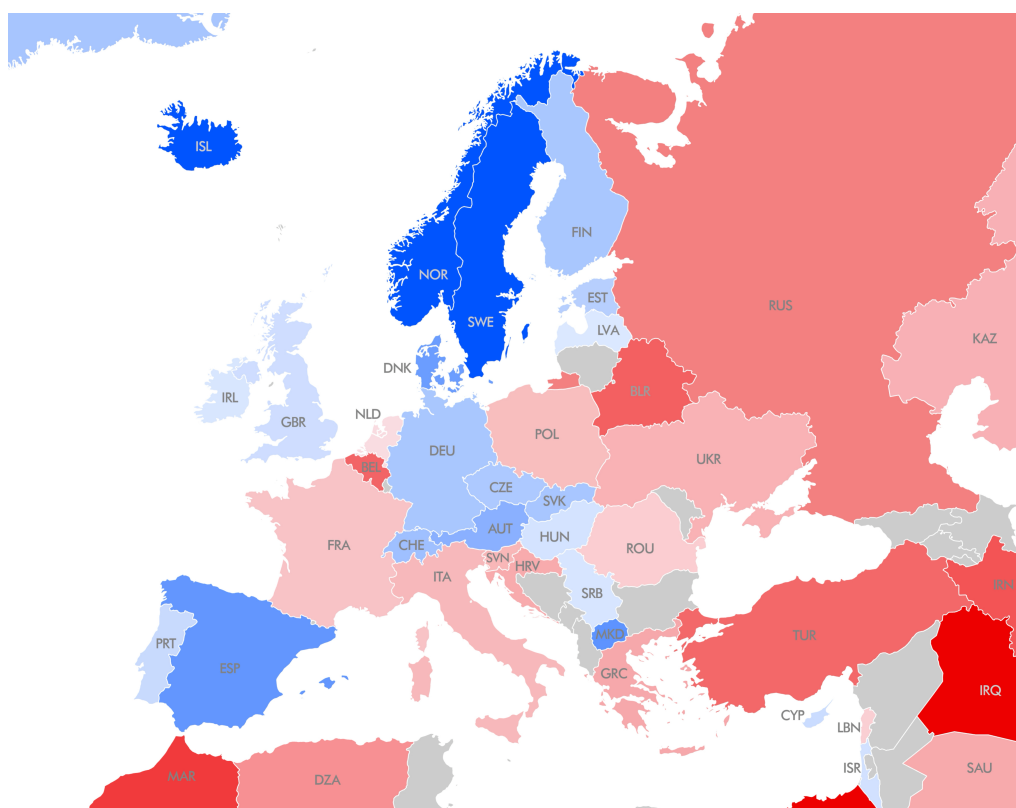
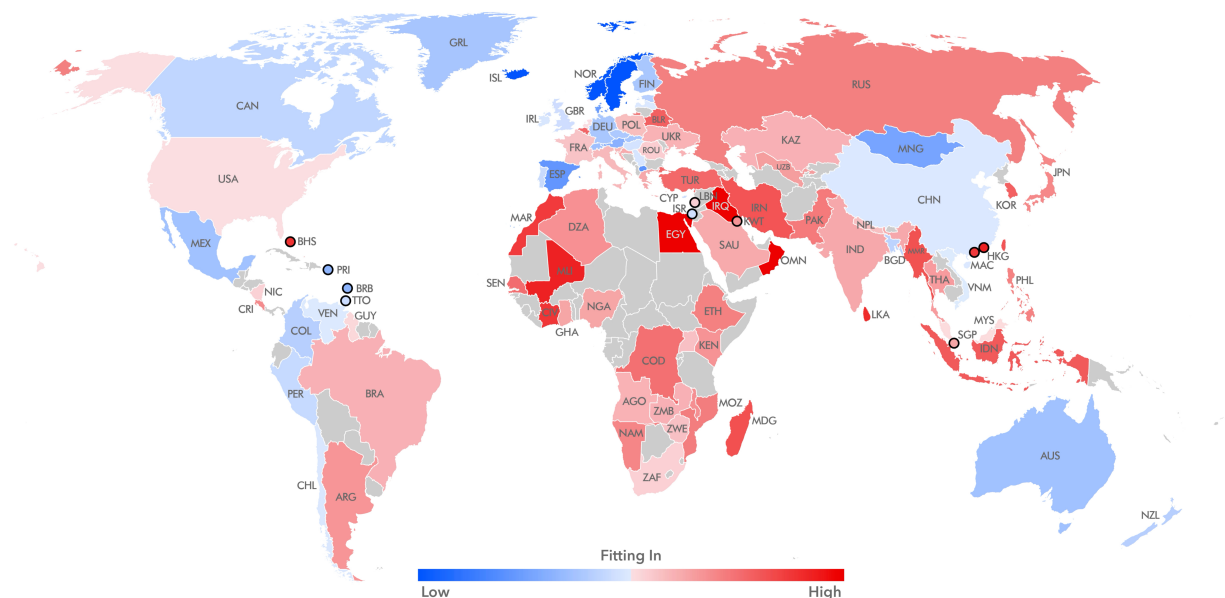
**Fig. S7. Responsibilism Component Scores for Guarding Personal Details.** Cultures in red scored higher on guarding personal details, which is more common in responsibilistic cultures. Scenarios in this component describe openly sharing personal information with strangers or people in the same group that you don't know well, such as the item, "You are working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You've been working on a sales account that could be a big accomplishment. The co-worker asks about the account. I would feel comfortable telling this co-worker about the account." This component is calculated from questions 4, 5, and 10 ([Table S4](#)).



**Fig. S8. Responsibilism Component Scores for Tight Friendship.** Cultures in red scored higher on tight friendship, which is more common in responsibilistic cultures. Items on this component have to do with making new friends versus sticking with friends, even if they are harsh to you, such as: “You get a 60% on a midterm, and you might fail the class if you don’t do better on the final exam. You’re in a bad mood. Your good friend criticizes you. ‘It’s your fault. You need to study harder.’ This makes you feel worse. This friend is being a good friend.” This component is calculated from questions 9, 11, and 13 ([Table S4](#)).

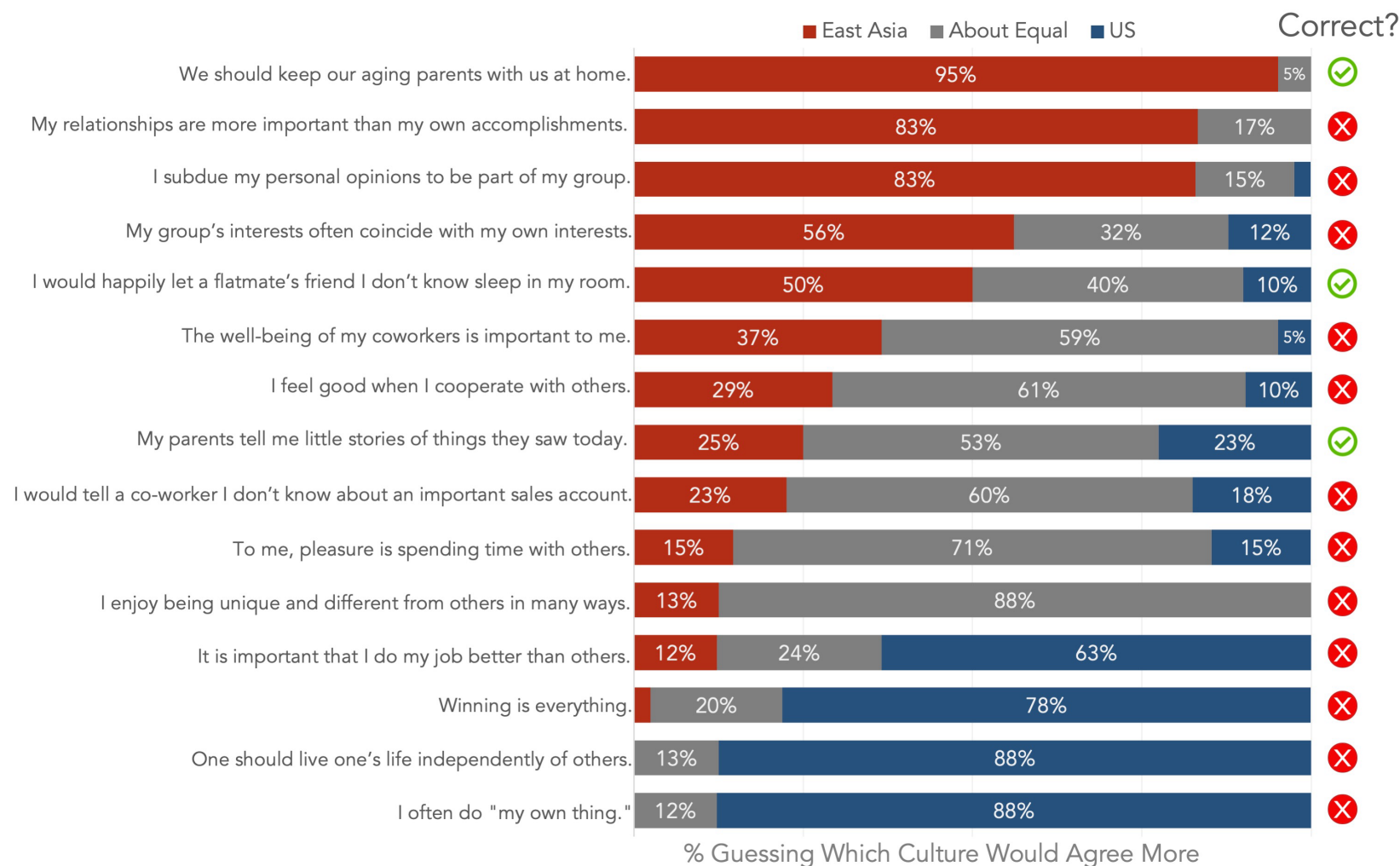


**Fig. S9. Responsibilism Component Scores for Stranger Danger.** Cultures in red scored higher on stranger danger, which is more common in responsibilistic cultures. Survey items on this component have to do with trusting strangers, such as the item, “You’re walking on the street, and a stranger starts talking to you. The stranger is probably trying to trick me.” This component is based on questions 2 and 3 ([Table S4](#)).

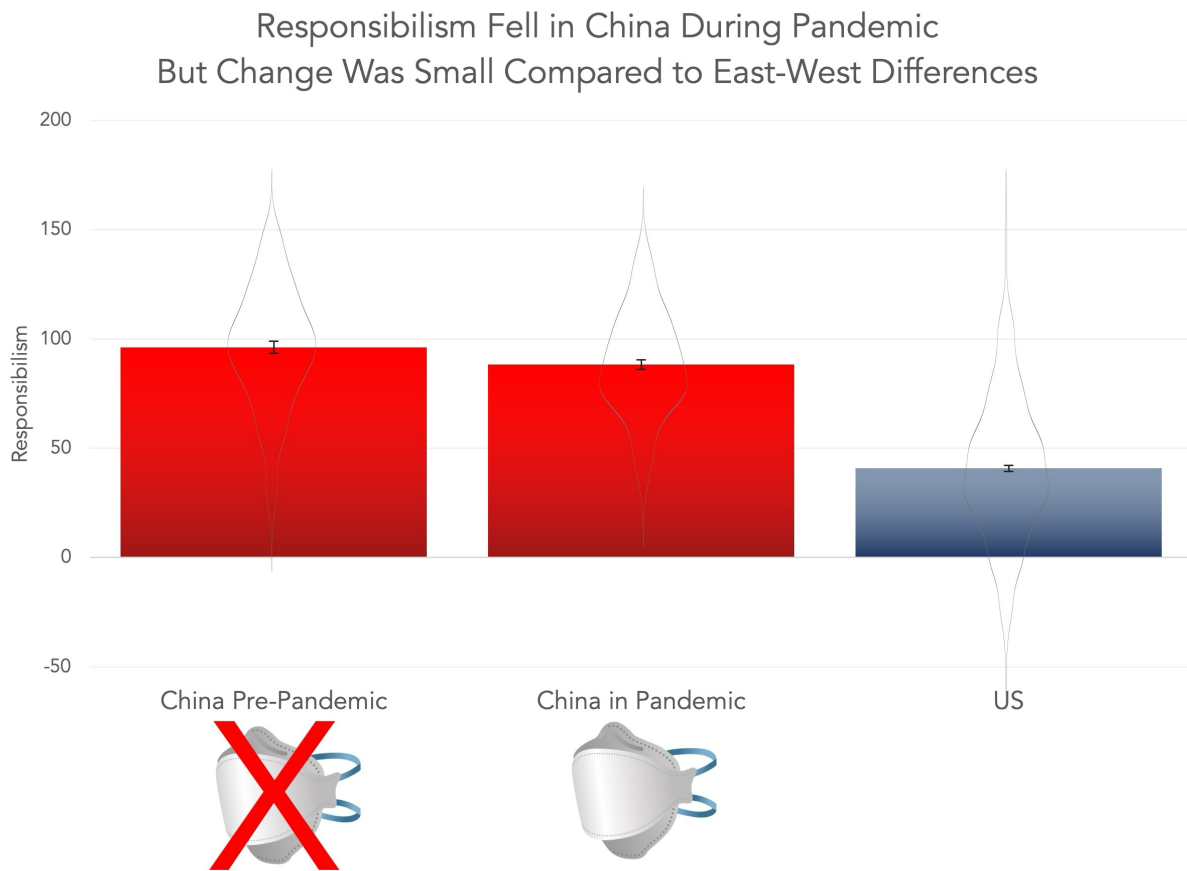


**Fig. S10. Responsibilism Component Scores for Fitting In.** Cultures in red scored higher on fitting in, which is more common in responsibilistic cultures. Scenarios on this component have to do with being bold in communicating with other people versus going along with the group. For example, one item describes asking for a raise: “Imagine you’ve worked for a company—you like your job, and you like your colleagues. But you haven’t gotten a raise for 5 years. Now you’re thinking about asking your boss for a raise. Of course, it could anger your boss and make working there uncomfortable. I would still ask my boss for a raise.” This component is calculated from questions 7, 12, and 14 ([Table S4](#)).

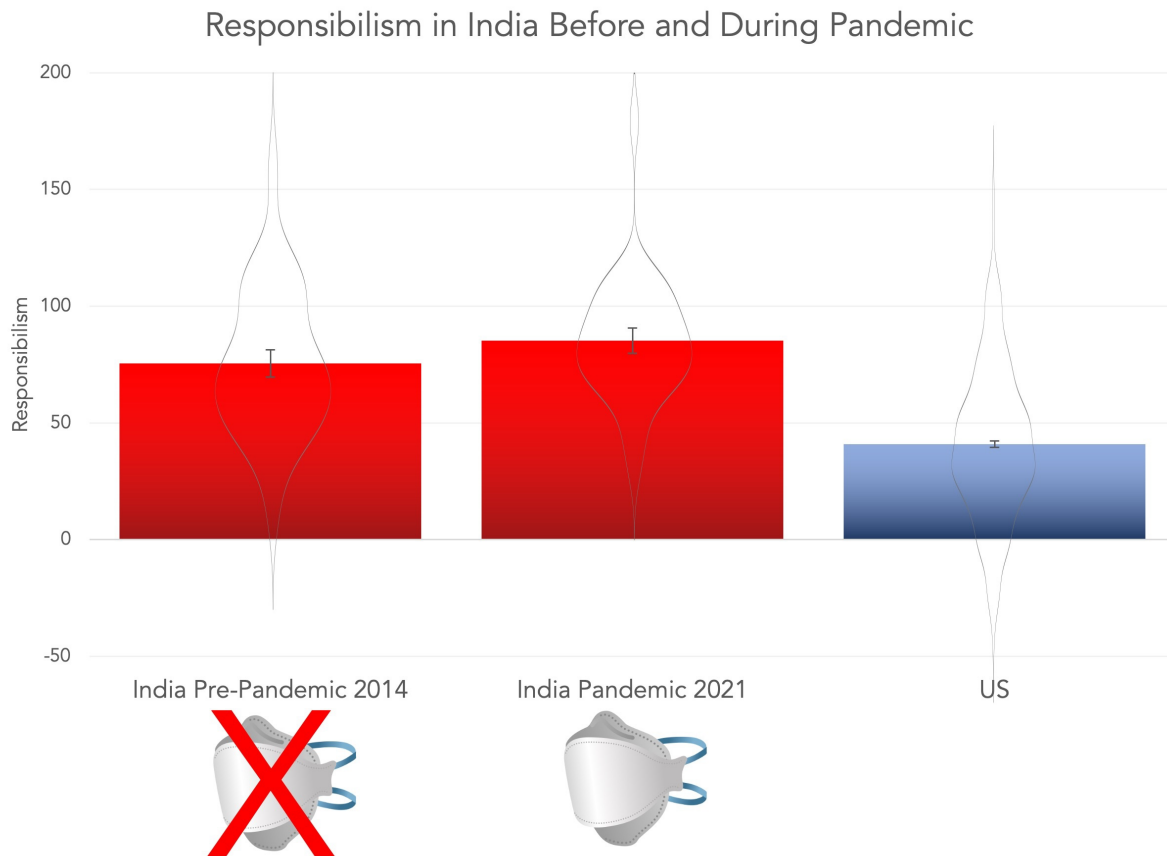
## Cultural Psychologists' Guesses



**Fig. S11. Researchers' Guesses About Cultural Differences Versus Actual Cultural Differences.** We coded guesses as correct if the most popular guess matched statistical significance. We coded "about equal" as correct if there was no significant difference. We abbreviated some question wordings for space. Full wordings are in [Table S13](#). [Table S14](#) reports demographics of the 40 cultural researchers surveyed. [Tables S13A](#) and [S13B](#) report the actual survey results from the US and China that we compared the guesses to.



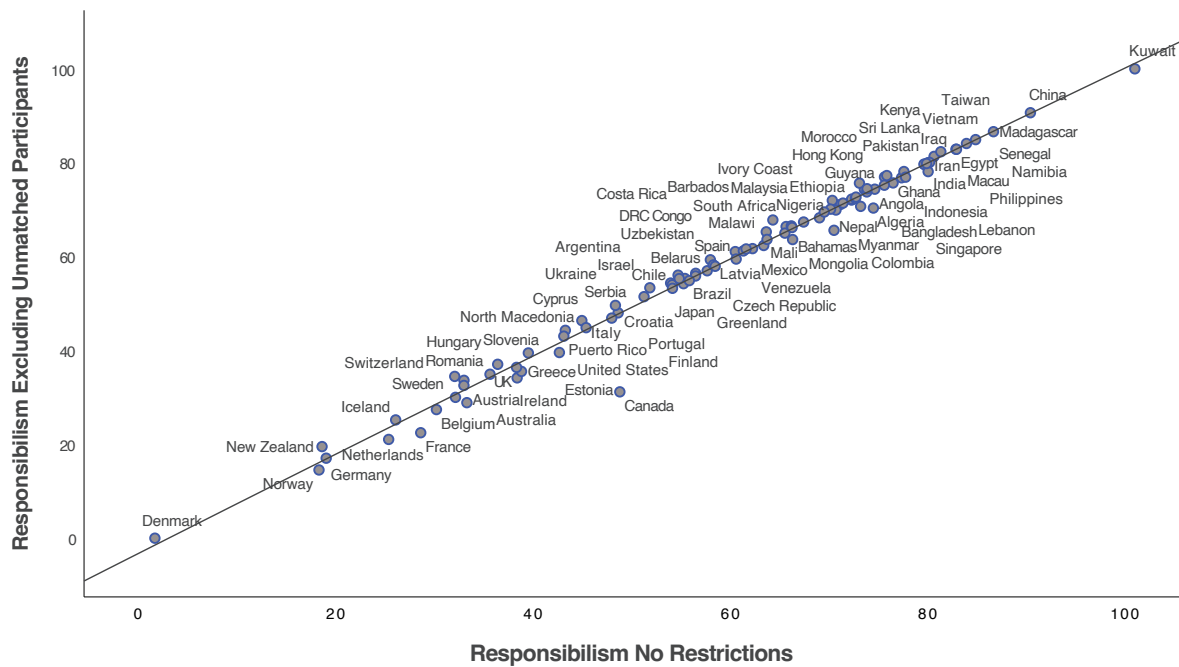
**Fig. S12. Responsibilism Went Down in China During the COVID-19 Pandemic (Left) But the Change Was Far Smaller Than East-West Differences (Right).** Data from China comes from a longitudinal sample of 104 college students at Hubei University in Wuhan. Wuhan was the epicenter of the COVID-19 outbreak in China. Students completed the responsibilism scale in 2019 and again in 2020. Scores can be below 0 because 0 corresponds to the lowest average score for a culture (Denmark). Of course, an individual can score below the average for Denmark.  $N = 104$  longitudinal participants in China and 511 in the US.



**Fig. S13. Responsibilism Went Up in India from 2014 Before the COVID-19 Pandemic (Left) to 2021 (Middle), But Both Samples Showed Higher Responsibilism Than US (Right).**

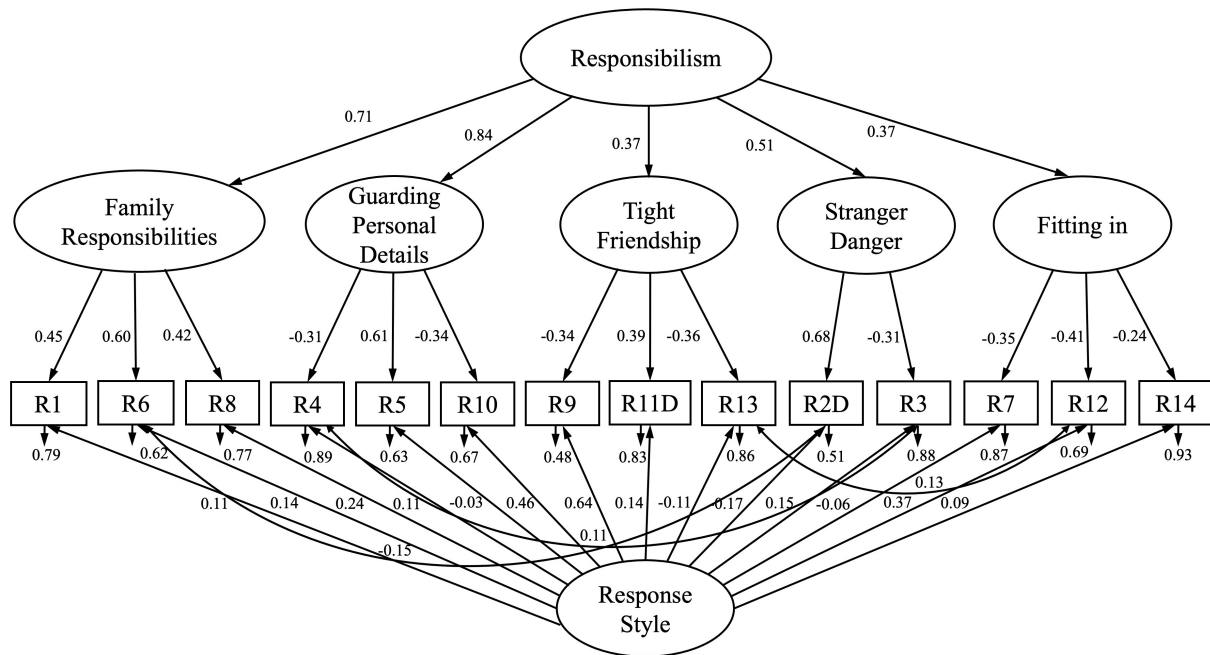
Data from India comes from a longitudinal sample of 28 participants on MTurk in October, 2014 and late April, 2021. The second data collection happened during India's intense delta wave, which overwhelmed hospitals and [killed 240,000 people](#). Bars = 1 SEM. Scores can be below 0 because 0 corresponds to the lowest average score for a culture (Denmark). Of course, an individual can score below the average for Denmark. Readers should be aware that the longitudinal sample size in India is small. It has 80% statistical power to detect effects of  $d = 0.55$ .



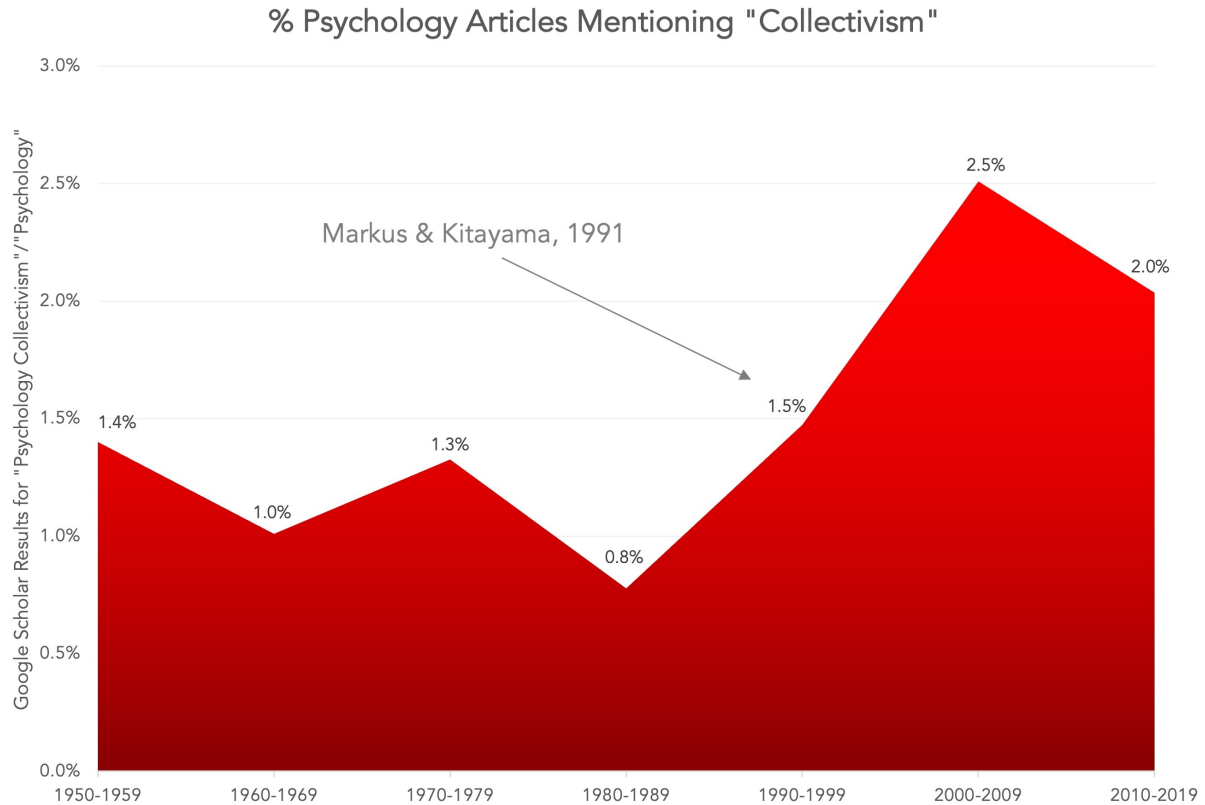


**Fig. S14. Responsibilism Scores with No Sample Restrictions (X Axis) Versus Excluding Participants with Non-Local Cultural Backgrounds (Y Axis).** Scores on the Y axis are after excluding participants who were born or grew up in a different culture and participants who reported a cultural identity different from the local culture (which we call “unmatched” participants). For example, this would exclude an American student studying abroad in Japan and taking the survey there. It would also exclude a participant in Egypt who reports a cultural background from Saudi Arabia. Although we think these restrictions are important, the restrictions make little difference to the results. The two calculations are correlated highly,  $r(98) = 0.99$ ,  $P < 0.001$ , 95% CI [0.990, 0.996].  $N = 100$  cultures.

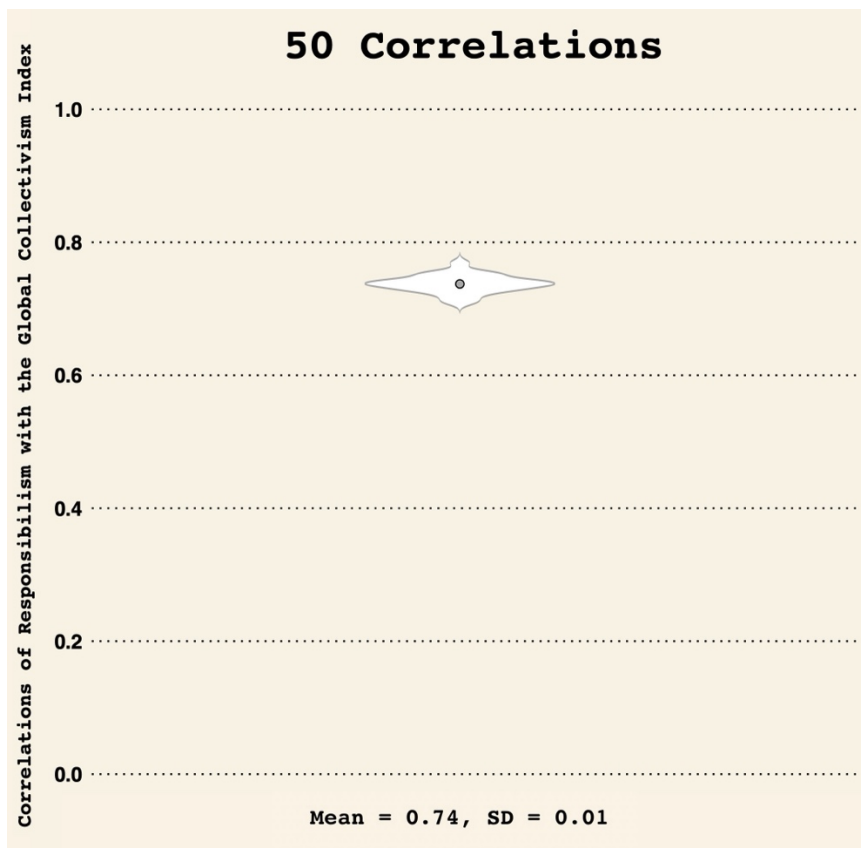




**Fig. S16. Multigroup factor analysis.** Model results for a multigroup factor analysis. This model tests the factor structure of the responsibilism scale using both individual-level responses while taking cultural groups into account. This analysis uses raw survey responses but adds a factor (bottom) to account for response styles (sometimes called a “common method factor”). We ran the model in the program MPlus. Loadings are standardized. The scale wordings and numbers are in [Table S4](#). Items with “D” are paired items, where one question asks about a responsibilistic response to a scenario, and the other item asks about a non-responsibilistic response to the scenario (such as item 2). The model analyzes paired items as a difference score with answers to the responsibilistic question minus answers to the non-responsibilistic question. The model had fit statistics of RMSEA = 0.021 vs. a suggested cutoff of < 0.08<sup>107</sup>, CFI = 0.908 vs. a suggested cutoff of > 0.90<sup>108,109</sup>, and SRMR = 0.032 vs. a suggested cutoff of < 0.08<sup>107</sup>. Because the “fitting in” facet had one of the lowest factor loadings, we tested whether removing this facet would change the performance of responsibilism culture scores. Removing the “fitting in” facet slightly decreased external validity performance from  $r = 0.675$  (all items) to  $r = 0.653$  (removing the three “fitting in” items). Similarly, removing these items slightly decreased sensitivity to cultural differences. Removing the “fitting in” facet decreased the correlation of responsibilism scores with the West, “sometimes West,” and non-West variable from  $r = 0.866$  to  $r = 0.841$ .  $N = 12,247$  participants.



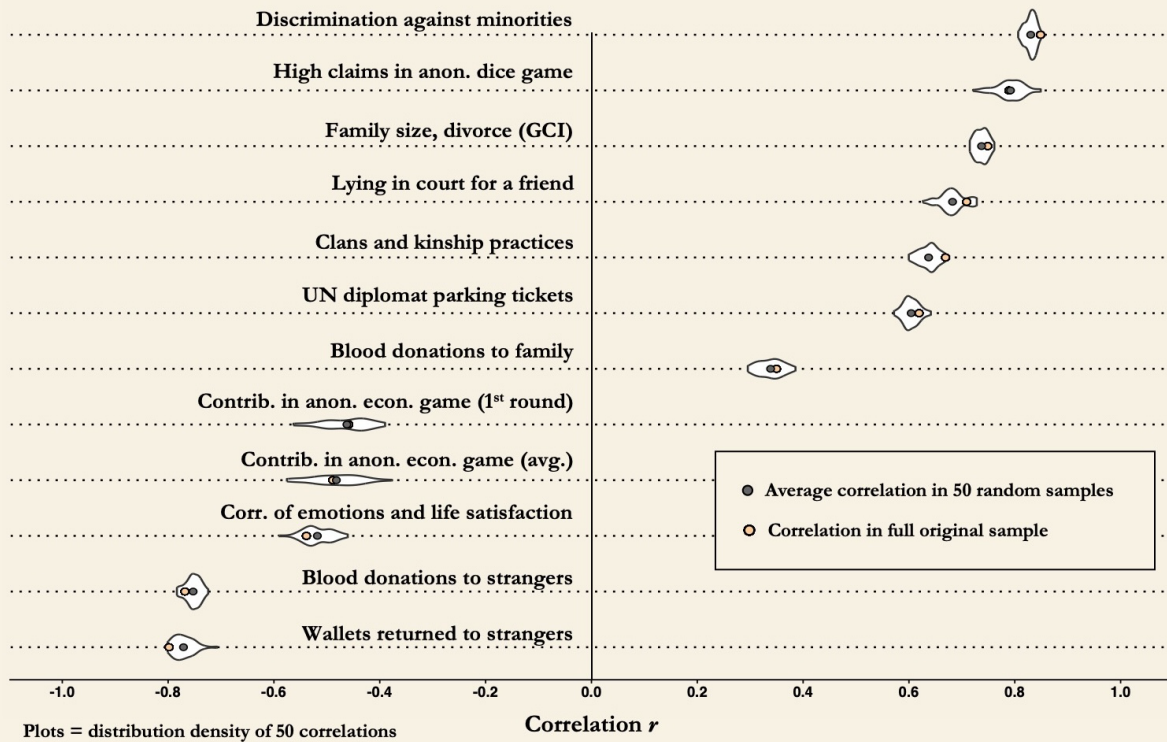
**Fig. S17. Percentage of Psychology Articles Mentioning “Collectivism” from 1950 to 2019.** This graph presents the number of Google Scholar search results for each decade mentioning “collectivism” and “psychology” divided by the number of articles mentioning “psychology.” The results show a spike around the time of Markus and Kitayama’s article<sup>1</sup>. We collected this data on November 5, 2024.



**Fig. S18A. Distribution of Correlations of Responsibility with the Global Collectivism Index in 50 Random Samples of 40 Participants per Culture.**

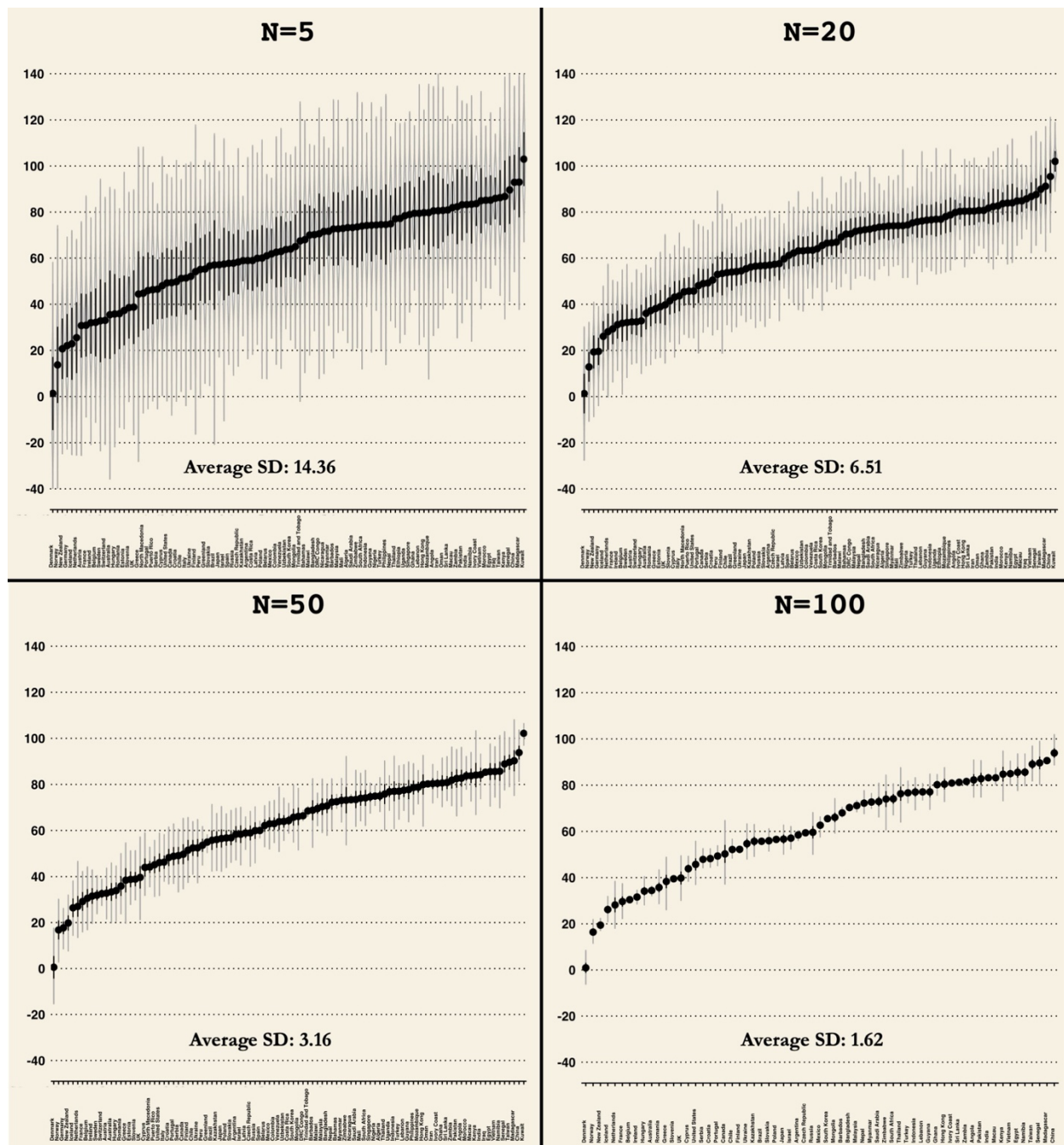
This violin plot shows the distribution of correlations between responsibility and the Global Collectivism Index in 50 random samples. Each estimate uses a random sample of 40 participants within each culture. The dot shows the average correlation. The width of the plot shows how many estimates were at that value. The results show that the correlation between responsibility and the Global Collectivism Index was fairly stable (falling within a narrow range), even with 40 participants per culture. This suggests that sample sizes of 40 participants per culture are sufficient (even if not ideal) to estimate responsibility scores that demonstrate external validity.

## External Validity Correlations in 50 Random Samples of $N = 40$ Per Culture



**Fig. S18B. Distribution of External Validity Correlations from 50 Random Samples of 40 Participants per Culture.** This graph shows the distribution of correlations between

responsibilism and markers of external validity (listed in Extended Data Tables 1-2). We randomly sampled 40 participants from each culture (from the full sample before any data-quality screening). We then calculated responsibilism scores for each culture and correlated these scores with external markers of collectivism, such as living with extended family. We then repeated this process 50 times. The results are similar to the correlations in the full sample (Fig. 5), which are printed on here with peach-colored circles. The results suggest that the correlations are stable across random samples and robust to forcing the sample sizes to be equal for all cultures. The four widest distributions are the markers of external validity that have the smallest number of cultures (such as the economic games, which have data for just 15 cultures).



**Fig. S19. Estimates of Culture Scores for Responsibilism Across 50 Random Samples Ranging from 5 to 100 Participants per Culture.** Each panel shows the spread of estimates for culture scores for responsibilism across 50 random samples within the larger dataset. For example, the top-left panel shows 50 random samples of 5 participants per culture. Each dot represents the average for a culture. The gray lines extend to the most extreme estimates. The black lines extend  $\pm 1$  SD from the average. In normal distributions, approximately 68% of all estimates should fall within the black lines. With samples of 100 (lower right), the black lines are so small that they are not visible.



**Table S1. Responsibilism predicts external markers of collectivism more strongly than previous collectivism scales.**

| Marker                                                                                      | Responsibilism |                       |                | Collectivism<br>(Singelis et al., 1995) |                       |                | Interdependent<br>Self-Construal<br>(Gudykunst et al., 1996) |                       |                | Collectivism<br>(Earley, 1993) |                       |                | Relational-Int.<br>Self-Construal<br>(Cross et al., 2000) |                       |                | Interdependent Self-<br>Construal<br>(Singelis, 1994) |                       |               |
|---------------------------------------------------------------------------------------------|----------------|-----------------------|----------------|-----------------------------------------|-----------------------|----------------|--------------------------------------------------------------|-----------------------|----------------|--------------------------------|-----------------------|----------------|-----------------------------------------------------------|-----------------------|----------------|-------------------------------------------------------|-----------------------|---------------|
|                                                                                             | <i>r</i>       | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                                | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                                                     | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                       | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                                                  | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                                              | <i>P</i> <sup>b</sup> | 95% CI        |
| <b>Tight Ties in Close Relationships</b>                                                    |                |                       |                |                                         |                       |                |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| Global Collectivism Index                                                                   | 0.85           | < 0.001               | [0.76, 0.92]   | 0.57                                    | < 0.001               | [0.34, 0.73]   | -0.68                                                        | < 0.001               | [-0.81, -0.50] | -0.23                          | 0.230                 | [-0.47, 0.06]  | -0.37                                                     | 0.072                 | [-0.59, -0.11] | 0.17                                                  | 0.816                 | [-0.18, 0.47] |
| Kinship Intensity Index                                                                     | 0.72           | < 0.001               | [0.55, 0.83]   | 0.42                                    | 0.006                 | [0.17, 0.63]   | -0.70                                                        | < 0.001               | [-0.82, -0.53] | -0.29                          | 0.129                 | [-0.52, -0.01] | -0.35                                                     | 0.072                 | [-0.57, -0.08] | -0.06                                                 | 0.899                 | [-0.39, 0.28] |
| Blood Donations to Family                                                                   | 0.39           | 0.007                 | [0.12, 0.60]   | 0.20                                    | 0.244                 | [-0.08, 0.46]  | -0.29                                                        | 0.060                 | [-0.53, -0.01] | 0.08                           | 0.575                 | [-0.20, 0.36]  | -0.16                                                     | 0.532                 | [-0.42, 0.13]  | 0.19                                                  | 0.816                 | [-0.15, 0.50] |
| Lying for Friend in Court                                                                   | 0.69           | < 0.001               | [0.44, 0.84]   | 0.03                                    | 0.859                 | [-0.33, 0.39]  | -0.51                                                        | 0.007                 | [-0.74, -0.19] | -0.35                          | 0.144                 | [-0.63, 0.01]  | -0.09                                                     | 0.892                 | [-0.44, 0.28]  | 0.05                                                  | 0.899                 | [-0.35, 0.43] |
| <b>Weak Responsibilities to Strangers and Anonymous Systems</b>                             |                |                       |                |                                         |                       |                |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| UN Diplomat Parking Tickets                                                                 | 0.65           | < 0.001               | [0.46, 0.79]   | 0.44                                    | 0.004                 | [0.19, 0.65]   | -0.50                                                        | < 0.001               | [-0.68, -0.25] | -0.13                          | 0.470                 | [-0.39, 0.16]  | -0.21                                                     | 0.377                 | [-0.46, 0.08]  | 0.41                                                  | 0.204                 | [0.08, 0.66]  |
| Dishonest Claims Anon. Dice Game <sup>a</sup>                                               | 0.77           | 0.002                 | [0.43, 0.92]   | 0.08                                    | 0.843                 | [-0.45, 0.57]  | -0.58                                                        | 0.036                 | [-0.84, -0.09] | -0.57                          | 0.108                 | [-0.84, -0.08] | -0.11                                                     | 0.892                 | [-0.59, 0.42]  | -0.09                                                 | 0.899                 | [-0.68, 0.57] |
| Discrimination Against Outgroups                                                            | 0.91           | < 0.001               | [0.84, 0.95]   | 0.48                                    | < 0.001               | [0.23, 0.67]   | -0.70                                                        | < 0.001               | [-0.82, -0.52] | -0.34                          | 0.102                 | [-0.56, -0.06] | -0.33                                                     | 0.080                 | [-0.56, -0.06] | 0.12                                                  | 0.816                 | [-0.22, 0.44] |
| <b>Pro-Sociality with Strangers</b>                                                         |                |                       |                |                                         |                       |                |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| Blood Donations Strangers                                                                   | -0.75          | < 0.001               | [-0.85, -0.59] | -0.42                                   | 0.007                 | [-0.63, -0.15] | 0.54                                                         | < 0.001               | [0.30, 0.71]   | 0.19                           | 0.353                 | [-0.10, 0.45]  | 0.21                                                      | 0.377                 | [-0.08, 0.47]  | -0.16                                                 | 0.816                 | [-0.47, 0.19] |
| Wallet Return Rate to Strangers <sup>a</sup>                                                | -0.81          | < 0.001               | [-0.90, -0.64] | -0.31                                   | 0.162                 | [-0.60, 0.04]  | 0.54                                                         | 0.004                 | [0.23, 0.75]   | 0.44                           | 0.102                 | [0.11, 0.68]   | 0.16                                                      | 0.651                 | [-0.20, 0.48]  | -0.17                                                 | 0.816                 | [-0.54, 0.27] |
| Contrib. in Anon. Econ. Game <sup>a</sup>                                                   | -0.53          | 0.099                 | [-0.86, 0.11]  | -0.30                                   | 0.489                 | [-0.76, 0.36]  | 0.50                                                         | 0.124                 | [-0.14, 0.85]  | 0.27                           | 0.470                 | [-0.40, 0.75]  | 0.07                                                      | 0.892                 | [-0.56, 0.64]  | -0.14                                                 | 0.899                 | [-0.71, 0.54] |
| Contrib....1 <sup>st</sup> Round <sup>a</sup>                                               | -0.52          | 0.099                 | [-0.86, 0.11]  | -0.22                                   | 0.619                 | [-0.72, 0.44]  | 0.41                                                         | 0.213                 | [-0.25, 0.81]  | 0.28                           | 0.470                 | [-0.38, 0.76]  | -0.05                                                     | 0.892                 | [-0.63, 0.57]  | 0.26                                                  | 0.816                 | [-0.44, 0.77] |
| <b>Basing Life Satisfaction on Internal Feelings (as Opposed to Social Roles or Status)</b> |                |                       |                |                                         |                       |                |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| Life Sat. Correlation w/ Emotion <sup>a</sup>                                               | -0.53          | 0.005                 | [-0.75, -0.19] | -0.32                                   | 0.170                 | [-0.62, 0.06]  | 0.34                                                         | 0.095                 | [-0.04, 0.63]  | 0.18                           | 0.470                 | [-0.21, 0.52]  | -0.06                                                     | 0.892                 | [-0.43, 0.32]  | -0.02                                                 | 0.920                 | [-0.45, 0.41] |
| Average Correlation                                                                         | 0.71           |                       |                | 0.37                                    |                       |                | -0.56                                                        |                       |                | -0.21                          |                       |                | -0.25                                                     |                       |                | 0.10                                                  |                       |               |

Note: In the bottom row, we calculated correlations in the correct theoretical direction as positive (shaded in green, such as when higher collectivism scores predicted more kinship intensity) and correlations in the wrong direction as negative (shaded in red, such as when higher collectivism scores predicted *less* kinship intensity). <sup>a</sup>To calculate average correlations, we excluded these markers because they had data for fewer than 40 cultures. <sup>b</sup>All *P* values in this table are adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S2. Correlations of collectivism scales and external markers of collectivism controlling for economic development.**

| Marker                                                                                      | Responsibilism |                       |                | Collectivism<br>(Singelis et al., 1995) |                       |               | Interdependent<br>Self-Construal<br>(Gudykunst et al., 1996) |                       |                | Collectivism<br>(Earley, 1993) |                       |                | Relational-Int.<br>Self-Construal<br>(Cross et al., 2000) |                       |                | Interdependent Self-<br>Construal<br>(Singelis, 1994) |                       |               |
|---------------------------------------------------------------------------------------------|----------------|-----------------------|----------------|-----------------------------------------|-----------------------|---------------|--------------------------------------------------------------|-----------------------|----------------|--------------------------------|-----------------------|----------------|-----------------------------------------------------------|-----------------------|----------------|-------------------------------------------------------|-----------------------|---------------|
|                                                                                             | <i>r</i>       | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                                | <i>P</i> <sup>b</sup> | 95% CI        | <i>r</i>                                                     | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                       | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                                                  | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>                                              | <i>P</i> <sup>b</sup> | 95% CI        |
| <b>Tight Ties in Close Relationships</b>                                                    |                |                       |                |                                         |                       |               |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| Global Collectivism Index                                                                   | 0.61           | < 0.001               | [0.39, 0.77]   | 0.26                                    | 0.320                 | [-0.14, 0.52] | -0.46                                                        | 0.004                 | [-0.69, -0.13] | -0.24                          | 0.226                 | [-0.46, -0.03] | -0.28                                                     | 0.534                 | [-0.54, 0.01]  | 0.17                                                  | 0.805                 | [-0.22, 0.45] |
| Kinship Intensity Index                                                                     | 0.50           | < 0.001               | [0.28, 0.66]   | 0.14                                    | 0.459                 | [-0.19, 0.48] | -0.55                                                        | < 0.001               | [-0.73, -0.26] | -0.28                          | 0.216                 | [-0.49, -0.04] | -0.24                                                     | 0.534                 | [-0.47, 0.001] | -0.24                                                 | 0.805                 | [-0.55, 0.11] |
| Blood Donations to Family                                                                   | 0.01           | 0.972                 | [-0.28, 0.25]  | -0.09                                   | 0.572                 | [-0.39, 0.20] | -0.003                                                       | 0.996                 | [-0.32, 0.35]  | 0.16                           | 0.483                 | [-0.05, 0.38]  | -0.04                                                     | 0.812                 | [-0.33, 0.27]  | 0.16                                                  | 0.805                 | [-0.18, 0.48] |
| Lying for Friend in Court                                                                   | 0.33           | 0.118                 | [-0.18, 0.65]  | -0.32                                   | 0.320                 | [-0.63, 0.19] | -0.25                                                        | 0.384                 | [-0.60, 0.20]  | -0.17                          | 0.488                 | [-0.61, 0.38]  | -0.08                                                     | 0.812                 | [-0.40, 0.32]  | -0.03                                                 | 0.898                 | [-0.47, 0.48] |
| <b>Weak Responsibilities to Strangers and Anonymous Systems</b>                             |                |                       |                |                                         |                       |               |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| UN Diplomat Parking Tickets                                                                 | 0.36           | 0.026                 | [0.04, 0.60]   | 0.16                                    | 0.459                 | [-0.21, 0.45] | -0.23                                                        | 0.354                 | [-0.51, 0.14]  | -0.08                          | 0.640                 | [-0.35, 0.19]  | -0.06                                                     | 0.812                 | [-0.31, 0.23]  | 0.44                                                  | 0.132                 | [0.16, 0.67]  |
| Dishonest Claims Anon. Dice Game <sup>a</sup>                                               | 0.51           | 0.093                 | [0.01, 0.82]   | -0.39                                   | 0.367                 | [-0.68, 0.28] | -0.21                                                        | 0.619                 | [-0.63, 0.42]  | -0.48                          | 0.216                 | [-0.76, 0.10]  | 0.12                                                      | 0.812                 | [-0.47, 0.80]  | -0.22                                                 | 0.805                 | [-0.89, 0.79] |
| Discrimination Against Outgroups                                                            | 0.78           | < 0.001               | [0.67, 0.86]   | 0.14                                    | 0.459                 | [-0.33, 0.51] | -0.50                                                        | < 0.001               | [-0.72, -0.19] | -0.38                          | 0.084                 | [-0.59, -0.17] | -0.20                                                     | 0.684                 | [-0.44, 0.07]  | 0.07                                                  | 0.805                 | [-0.42, 0.46] |
| <b>Pro-Sociality with Strangers</b>                                                         |                |                       |                |                                         |                       |               |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| Blood Donations Strangers                                                                   | -0.39          | 0.018                 | [-0.59, -0.06] | -0.03                                   | 0.853                 | [-0.38, 0.36] | 0.19                                                         | 0.384                 | [-0.16, 0.50]  | 0.16                           | 0.483                 | [-0.09, 0.39]  | 0.04                                                      | 0.812                 | [-0.23, 0.31]  | -0.13                                                 | 0.805                 | [-0.44, 0.27] |
| Wallet Return Rate to Strangers <sup>a</sup>                                                | -0.39          | 0.052                 | [-0.69, 0.01]  | 0.23                                    | 0.428                 | [-0.28, 0.53] | 0.18                                                         | 0.574                 | [-0.24, 0.53]  | 0.33                           | 0.216                 | [-0.14, 0.64]  | -0.06                                                     | 0.812                 | [-0.40, 0.28]  | -0.09                                                 | 0.805                 | [-0.48, 0.35] |
| Contrib. in Anon. Econ. Game <sup>a</sup>                                                   | -0.18          | 0.720                 | [-0.87, 0.74]  | -0.59                                   | 0.320                 | [-0.95, 0.85] | 0.25                                                         | 0.619                 | [-0.74, 0.97]  | -0.31                          | 0.488                 | [-0.88, 0.71]  | -0.14                                                     | 0.812                 | [-0.87, 0.85]  | -0.26                                                 | 0.805                 | [-0.96, 0.57] |
| Contrib....1 <sup>st</sup> Round <sup>a</sup>                                               | -0.18          | 0.720                 | [-0.83, 0.73]  | -0.48                                   | 0.367                 | [-0.87, 0.61] | 0.11                                                         | 0.910                 | [-0.74, 0.81]  | -0.27                          | 0.508                 | [-0.74, 0.73]  | -0.27                                                     | 0.812                 | [-0.80, 0.48]  | 0.24                                                  | 0.805                 | [-0.62, 0.87] |
| <b>Basing Life Satisfaction on Internal Feelings (as Opposed to Social Roles or Status)</b> |                |                       |                |                                         |                       |               |                                                              |                       |                |                                |                       |                |                                                           |                       |                |                                                       |                       |               |
| Life Sat. Correlation w/ Emotion <sup>a</sup>                                               | -0.06          | 0.841                 | [-0.50, 0.29]  | -0.15                                   | 0.550                 | [-0.48, 0.42] | 0.001                                                        | 0.996                 | [-0.39, 0.42]  | -0.07                          | 0.723                 | [-0.46, 0.43]  | -0.06                                                     | 0.812                 | [-0.36, 0.23]  | 0.08                                                  | 0.805                 | [-0.31, 0.50] |
| Average Correlation                                                                         | 0.43           |                       |                | 0.05                                    |                       |               | -0.31                                                        |                       |                | -0.16                          |                       |                | -0.13                                                     |                       |                | 0.10                                                  |                       |               |

Note: These correlations control for 2019 log GNI per capita. The bottom row calculates correlations in the correct theoretical direction as positive (shaded in green, such as higher collectivism scores predicting more kinship intensity) and correlations in the wrong direction as negative (shaded in red, such as when higher collectivism scores predicted *less* kinship intensity). <sup>a</sup>We excluded these markers from the averages because they had data for fewer than 40 cultures.

<sup>b</sup>All *P* values in this table are adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S3. Responsibilism predicts external markers of collectivism even controlling for previous scales.**

| Responsibilism Controlling for...                                                    |                       |                       |                |                        |                       |                |              |                       |                |                    |                       |                |                |                       |                |
|--------------------------------------------------------------------------------------|-----------------------|-----------------------|----------------|------------------------|-----------------------|----------------|--------------|-----------------------|----------------|--------------------|-----------------------|----------------|----------------|-----------------------|----------------|
| Marker                                                                               | Singelis et al., 1995 |                       |                | Gudykunst et al., 1996 |                       |                | Earley, 1993 |                       |                | Cross et al., 2000 |                       |                | Singelis, 1994 |                       |                |
|                                                                                      | <i>r</i>              | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>               | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>     | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>           | <i>P</i> <sup>b</sup> | 95% CI         | <i>r</i>       | <i>P</i> <sup>b</sup> | 95% CI         |
| Tight Ties in Close Relationships                                                    |                       |                       |                |                        |                       |                |              |                       |                |                    |                       |                |                |                       |                |
| Global Collectivism Index                                                            | 0.83                  | < 0.001               | [0.73, 0.90]   | 0.73                   | < 0.001               | [0.59, 0.84]   | 0.85         | < 0.001               | [0.77, 0.92]   | 0.83               | < 0.001               | [0.74, 0.91]   | 0.77           | < 0.001               | [0.63, 0.88]   |
| Kinship Intensity Index                                                              | 0.67                  | < 0.001               | [0.48, 0.79]   | 0.47                   | 0.002                 | [0.24, 0.66]   | 0.69         | < 0.001               | [0.53, 0.81]   | 0.68               | < 0.001               | [0.55, 0.79]   | 0.68           | < 0.001               | [0.52, 0.81]   |
| Blood Donations Family                                                               | 0.34                  | 0.018                 | [0.13, 0.54]   | 0.28                   | 0.066                 | [0.04, 0.51]   | 0.46         | 0.002                 | [0.23, 0.64]   | 0.36               | 0.013                 | [0.14, 0.54]   | 0.34           | 0.060                 | [0.06, 0.59]   |
| Lying for Friend in Court                                                            | 0.69                  | < 0.001               | [0.46, 0.86]   | 0.57                   | 0.002                 | [0.34, 0.78]   | 0.64         | < 0.001               | [0.47, 0.80]   | 0.69               | < 0.001               | [0.50, 0.85]   | 0.63           | 0.002                 | [0.43, 0.80]   |
| Weak Responsibilities to Strangers and Anonymous Systems                             |                       |                       |                |                        |                       |                |              |                       |                |                    |                       |                |                |                       |                |
| UN Diplomat Parking Tickets                                                          | 0.58                  | < 0.001               | [0.35, 0.75]   | 0.50                   | < 0.001               | [0.30, 0.67]   | 0.66         | < 0.001               | [0.48, 0.80]   | 0.64               | < 0.001               | [0.47, 0.77]   | 0.49           | 0.005                 | [0.21, 0.72]   |
| Dishonest Claims Anon. Dice Game <sup>a</sup>                                        | 0.77                  | 0.002                 | [0.37, 0.92]   | 0.63                   | 0.018                 | [0.24, 0.86]   | 0.64         | 0.013                 | [0.26, 0.88]   | 0.77               | 0.002                 | [0.47, 0.93]   | 0.84           | 0.003                 | [0.65, 0.99]   |
| Discrimination Against Outgroups                                                     | 0.89                  | < 0.001               | [0.81, 0.93]   | 0.82                   | < 0.001               | [0.72, 0.89]   | 0.89         | < 0.001               | [0.84, 0.93]   | 0.89               | < 0.001               | [0.84, 0.94]   | 0.90           | < 0.001               | [0.85, 0.94]   |
| Pro-Sociality with Strangers                                                         |                       |                       |                |                        |                       |                |              |                       |                |                    |                       |                |                |                       |                |
| Blood Donations Strangers                                                            | -0.71                 | < 0.001               | [-0.83, -0.53] | -0.62                  | < 0.001               | [-0.77, -0.45] | -0.75        | < 0.001               | [-0.84, -0.61] | -0.74              | < 0.001               | [-0.84, -0.58] | -0.67          | < 0.001               | [-0.83, -0.47] |
| Wallet Return Rate to Strangers <sup>a</sup>                                         | -0.79                 | < 0.001               | [-0.91, -0.55] | -0.72                  | < 0.001               | [-0.87, -0.50] | -0.76        | < 0.001               | [-0.88, -0.56] | -0.80              | < 0.001               | [-0.92, -0.63] | -0.75          | < 0.001               | [-0.90, -0.52] |
| Contrib. in Anon. Econ. Game <sup>a</sup>                                            | -0.56                 | 0.081                 | [-0.93, -0.11] | -0.20                  | 0.562                 | [-0.72, 0.50]  | -0.62        | 0.047                 | [-0.96, 0.07]  | -0.61              | 0.048                 | [-0.94, -0.11] | -0.35          | 0.345                 | [-0.90, 0.53]  |
| ...1 <sup>st</sup> Round <sup>a</sup>                                                | -0.54                 | 0.085                 | [-0.89, 0.20]  | -0.38                  | 0.269                 | [-0.88, 0.41]  | -0.58        | 0.059                 | [-0.90, 0.14]  | -0.69              | 0.021                 | [-0.92, -0.20] | -0.17          | 0.631                 | [-0.89, 0.27]  |
| Basing Life Satisfaction on Internal Feelings (as Opposed to Social Roles or Status) |                       |                       |                |                        |                       |                |              |                       |                |                    |                       |                |                |                       |                |
| Life Sat. Correlation w/ Emotion <sup>a</sup>                                        | -0.53                 | 0.005                 | [-0.78, -0.14] | -0.44                  | 0.028                 | [-0.72, -0.09] | -0.53        | 0.005                 | [-0.80, 0.03]  | -0.53              | 0.005                 | [-0.80, -0.19] | -0.61          | 0.005                 | [-0.88, -0.23] |
| Average Correlation                                                                  | 0.67                  |                       |                | 0.57                   |                       |                | 0.71         |                       |                | 0.69               |                       |                | 0.64           |                       |                |

Note: This table reports partial correlations between responsibilism and markers of external validity controlling for each prior collectivism scale. Correlations in the wrong theoretical direction are counted as negative in the calculation of average correlations. <sup>a</sup>To calculate average correlations, we excluded these markers because they had data for fewer than 40 cultures. <sup>b</sup>All *P* values in this table are adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S4. The US English version of the 14-item responsibilism scale.**

|     | Item                                                                                                                                                                                                                                                                                                                                 | Rationale for Creating the Item                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | External Validity Correlation |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|
| 1   | Children should be taught to place duty before pleasure.                                                                                                                                                                                                                                                                             | Responsibilistic cultures emphasize duties in close relationships. The use of “children” implies that this is within the family. This item comes from Singelis and colleagues’ 1995 scale <sup>3</sup> , which demonstrates that the earlier scales contained some items that tapped into the responsibilistic vision of collectivism.                                                                                                                                                                                                                                                     | $r = 0.37$                    |
| 2A  | You’re walking on the street, and a stranger starts talking to you.<br><b>The stranger is probably trying to trick me.</b>                                                                                                                                                                                                           | Responsibilistic cultures have stronger duties to close others but weaker duties and trust toward strangers <sup>36,110</sup> . This item avoids vague wordings <sup>10</sup> by specifying that the other person is a stranger and that the context is walking on the street.                                                                                                                                                                                                                                                                                                             | $r = 0.29$                    |
| 2B* | I would trust the stranger.<br>The person probably has good intentions.                                                                                                                                                                                                                                                              | Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items. This can reveal cultural differences in how people answer surveys <sup>20</sup> .                                                                                                                                                                                                                                                                                                                                                                                   |                               |
| 3*  | It’s after dinner, and you go to the ATM to get some cash. The person in front of you turns around and says, “I can’t figure out how to get the machine to take a deposit. Can you help me do it?”<br><b>I would trust the person.</b>                                                                                               | Responsibilistic cultures have stronger duties to close others but weaker duties and trust toward strangers <sup>36,110</sup> . This item avoids abstract, unclear wordings by specifying that the other person is a stranger and that the context is evening and at an ATM. We anticipate needing to update this item as cashless payments become more common.                                                                                                                                                                                                                            | $r = 0.39$                    |
| 4*  | You are working in a company for two years, and one day you meet a co-worker from a different department that you’ve never met before. You’ve been working on a sales account that could be a big accomplishment. The co-worker asks about the account.<br><b>I would feel comfortable telling this co-worker about the account.</b> | Responsibilistic cultures are more wary of non-close ties <sup>36,110</sup> . This item specifically tests the idea that belonging to the same big group is <i>not</i> sufficient for tight ties in responsibilistic cultures. In this scenario, the other person is in the same group (the same company) but is unfamiliar. We theorized that people in responsibilistic cultures would trust this co-worker less because the tie is weak. This fits with prior research finding that people in collectivistic cultures are more vigilant toward people in large ingroups <sup>27</sup> . | $r = 0.55$                    |

|   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                              |
|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|
| 5 | <p>You're working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You are sitting outside a meeting room waiting for the meeting to start, and the two of you start talking. The co-worker starts asking you if you're married, if you have children, what your hobbies are. Then you remember what a friend of yours told you: "Be careful not to share your personal secrets with people you don't know because they could use those secrets to hurt you."</p> <p><b>Do you agree with this statement?</b></p> | <p>People in responsibilistic cultures are more wary of non-close ties<sup>36</sup>. This item specifically tests the idea that belonging to the same big group is <i>not</i> sufficient for tight ties in responsibilistic cultures. In this scenario, the other person is in the same group (the same company) but is unfamiliar. We theorized that people in responsibilistic cultures would trust this co-worker less because the tie is weak. Studies on "enemyship" and vigilance have found that people in collectivistic cultures are more wary toward ingroup members in China and Ghana than the US<sup>27,64,103</sup>.</p> | <p><math>r = 0.52</math></p> |
|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|

|   |                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                              |                              |
|---|----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|
| 6 | <p>We should keep our aging parents with us at home.</p> | <p>Responsibilistic cultures emphasize duties in close relationships. The words "parents" and "home" imply helping in the family. This sort of help is pragmatic and objective. It does not ask about warm feelings. This item comes from Singelis and colleagues' 1995 scale<sup>3</sup>, which demonstrates that the earlier scales contained some items that tapped into the responsibilistic vision of collectivism.</p> | <p><math>r = 0.62</math></p> |
|---|----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|

|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 7* | <p>You start working at a new company. The salary is good, and you like the job. But when you get your first paycheck, you find out that you got more money than you expected. You look at the paycheck, and you see that they didn't take out the legally required taxes. You ask your boss, and your boss says, "Oh, everyone here does that. That's no problem. And this way you earn more money!"</p> <p><b>I would insist on paying the required taxes.</b></p> | <p>Responsibilistic cultures have stronger duties in close relationships but weaker ties in distant relationships and impersonal systems<sup>36,110</sup>. In the case of cheating on taxes, the harm is mostly to strangers—an impersonal system. The fact that everyone in the company evades this tax also suggests (a) that evading taxes would conform to the social norm and (b) breaking the rules would not likely bring punishment.</p> <p><math>r = 0.36</math></p> |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|   |                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 8 | <p>Children should feel honored if their parents receive a distinguished award.</p> | <p>Responsibilistic cultures emphasize duties in close relationships, which means that people's accomplishments are less individually determined. Thus, children may feel more closely tied to their parents' accomplishments and failures. The words "children" and "parents" specify close relationships. The description "distinguished award" implies social approval, similar to the concepts of honor and face, which tend to be stronger in collectivistic cultures<sup>11</sup>. This item comes from Singelis and colleagues' 1995 scale<sup>3</sup>, which is consistent with our argument that the earlier scales contained some items that tapped into the responsibilistic vision of collectivism.</p> <p><math>r = 0.53</math></p> |
|---|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|    |                                                                                                 |                                                                                                                                                                                                                                                                                                                                                                         |
|----|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 9* | <p>I like going to parties and talking to people I don't know in order to make new friends.</p> | <p>Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority<sup>20</sup>. It may even be strange to speak to strangers without some obvious connection to that person. If relationships come with strong duties, it may even be overly costly to build too many relationships.</p> <p><math>r = 0.18</math></p> |
|----|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|      |                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |            |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 10*  | <p>Imagine you're on a four-hour flight to another city, and you're traveling alone. Another person sits in the next seat to you.</p> <p><b>I would talk to that person and share details about myself—my job, my family members, where I went to school, whether I'm married or have a boyfriend/girlfriend.</b></p> | <p>Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority<sup>20</sup>. It may even be strange to speak to strangers without some obvious connection to that person. This item also taps into concerns about the potential for harm in building relationships too freely. The context of the plan and the details of the personal information keep the scenario concrete<sup>10</sup>. The details prevent people from filling in the details differently across cultures. In some cultures, people said that taking domestic flights was not common and suggested replacing this with a train or bus ride.</p> | $r = 0.42$ |
| 11A  | <p>You get a 60% on a midterm, and you might fail the class if you don't do better on the final exam. You're in a bad mood. Your good friend criticizes you. "It's your fault. You need to study harder." This makes you feel worse.</p> <p><b>This friend is being a good friend.</b></p>                            | <p>This item contrasts warm fuzzies versus more long-term duties about people's objective outcomes. In cultures where it is important to express warm feelings in a relationship, people should disagree with this item. But in cultures where people do not prioritize expressing warm feelings (and if they expect relationships to be binding enough that they can withstand harsh feedback), people should agree with this item.</p>                                                                                                                                                                                                                                  | $r = 0.18$ |
| 11B* | <p><b>I would try to make friends who are more supportive.</b></p>                                                                                                                                                                                                                                                    | <p>Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items (response style)<sup>17</sup>. Agreeing with this item implies that people believe it is possible/acceptable to leave old friends and make new friends, which means that relationships are not strongly binding.</p>                                                                                                                                                                                                                                                                                                                             |            |



|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 12* | <p>You're in a math class with 100 people in it. After class on Wednesday, you go the cafeteria to eat lunch. You sit at a table, and you are studying your notes from the class. The cafeteria is crowded, so another person sits down at your table. You recognize the person is in your class, but you've never talked to that person and don't know the person's name. You could continue studying your notes, or you could talk to that person "Hey, how do you like the class?"</p> <p><b>I would start a conversation with that person.</b></p> | <p>Responsibilistic cultures prioritize close relationships. Building relationships with weak ties is not usually a priority. This item specifies being in the same large group (the same school and the same class) but also that the other person is not a close tie. This is based on our theorizing that responsibilism is more about tight ties in specific relationships, not being in the same large group.</p> <p><math>r = 0.30</math></p> |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 13* | <p>One day, you discover your dream in life is to write novels. You spend two years writing and editing a novel. After it's ready, you share the novel with several of your friends. A few days later, they say, "This novel is silly! You should give up this dream. It's impractical. If you choose a practical job, you'll be happier." They make you so sad that you start to cry after you go back home.</p> <p><b>These people aren't really your friends. I should find friends that are more supportive.</b></p> | <p>This item contrasts warm fuzzies versus long-term duties about people's objective outcomes. If people think that relationships are binding and long lasting, they should expect that the friendship will last through the harsh feedback. If people think that expressing emotional support is a primary function of relationships, they should agree with this item. The original scenario described dating, but collaborators in some cultures said dating is taboo, so we edited the scenario to be about becoming a writer.</p> <p><math>r = 0.35</math></p> |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|     |                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |            |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 14* | <p>Imagine you've worked for a company—you like your job, and you like your colleagues. But you haven't gotten a raise for 5 years. Now you're thinking about asking your boss for a raise. Of course, it could anger your boss and make working there uncomfortable.</p> <p><b>I would still ask my boss for a raise.</b></p> | <p>This scenario asks about risking disrupting a work relationship. In cultures where relationships (even at work) tend to last longer, the consequences of upsetting the relationship will be harsher<sup>20</sup>. In cultures where relationships tend to be shorter, the costs are smaller. We adapted this item from Peng and colleagues, who argued that concrete scenarios are better at detecting cultural differences than abstract statements<sup>10</sup>.</p> | $r = 0.34$ |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|

Note: \*Items marked with asterisks are reverse coded so that higher values represent more responsibility. The response options were *Strongly disagree* (1), *Disagree* (2), *Slightly disagree* (3), *Neither agree nor disagree* (4), *Agree slightly* (5), *Agree* (6), *Strongly agree* (7). We coded the paired A-B questions as the responsibility question (no asterisk) minus the individualistic question (asterisk). For example, Item 2 is coded 2A – 2B. The external validity correlation is the average correlation of culture scores for each item with a basket of seven external indicators of collectivism, such as living with extended family and kinship practices. [Supplementary Section 15](#) details the seven indicators.

**Sub factors:**

**Family responsibilities:** Items 1, 6, 8

**Guarding personal details:** Items 4, 5, 10

**Tight friendship:** Items 9, 11A, 11B, 13

**Stranger danger:** Items 2A, 2B, 3

**Fitting in:** Items 7, 12, 14

**Table S5. The US English version the five-item responsibilism scale.**

|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Children should be taught to place duty before pleasure.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 2* | <p>You are working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You've been working on a sales account that could be a big accomplishment. The co-worker asks about the account.</p> <p><b>I would feel comfortable telling this co-worker about the account.</b></p>                                                                                                                                                                                                                                         |
| 3  | <p>You're working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You are sitting outside a meeting room waiting for the meeting to start, and the two of you start talking. The co-worker starts asking you if you're married, if you have children, what your hobbies are. Then you remember what a friend of yours told you: "Be careful not to share your personal secrets with people you don't know because they could use those secrets to hurt you."</p> <p><b>Do you agree with this statement?</b></p> |
| 4  | We should keep our aging parents with us at home.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 5  | Children should feel honored if their parents receive a distinguished award.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

Note: \*Reverse coded so that higher values represent more responsibilism. The response options were *Strongly disagree* (1), *Disagree* (2), *Slightly disagree* (3), *Neither agree nor disagree* (4), *Agree slightly* (5), *Agree* (6), *Strongly agree* (7). High reliability values should generally be more difficult to achieve with fewer scale items. The culture-level reliability of the entire scale is 0.89 ([Table S19A](#)) and 0.84 with the five-item version. Another limitation of the five-item scale is that it does not allow for the calculation of the sub-facets. Researchers who are interested in exploring some of the more specific components (such as tight friendship or caution around strangers) should use the full scale. Finally, the five-item version would not allow researchers to detect response style differences across cultures in detail. In this study, we use agreement across all items as a way of adjusting for response styles across cultures ([Supplementary Section 18](#)). With only five items, researchers will not be able to measure response style in as much detail as we did. Similarly, the five-item scale has only one reverse-coded item, which gives less ability to measure response style differences across positively and negatively coded items.

**Table S6. Sample demographics by culture (part 1).**

| <b>Culture</b> | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b> | <b>Site</b>        | <b>Compensation</b>      |
|----------------|----------|---------------|------------|-----------|-----------------|-----------------|--------------------|--------------------------|
| Algeria        | 65       | Stud./Adult   | 29.90      | 9.36      | 34.4            | Arabic          | Various            | Payment                  |
| Angola         | 204      | Student       | 25.76      | 5.87      | 52.9            | Portuguese      | Luanda             | Payment                  |
| Argentina      | 104      | Student       | 19.87      | 1.47      | 50.5            | Spanish         | Various            | Payment                  |
| Australia      | 103      | Student       | 24.47      | 8.47      | 82.8            | English         | Perth              | Course Credit            |
| Austria        | 90       | Stud./Adult   | 27.54      | 6.45      | 47.8            | German          | Various            | Payment                  |
| Bahamas        | 52       | Student       | 25.71      | 6.42      | 59.6            | English         | Various            | Prize Raffle             |
| Bangladesh     | 100      | Student       | 22.63      | 1.64      | 48.0            | Bengali         | Dhaka              | Payment, CC <sup>a</sup> |
| Barbados       | 107      | Student       | 28.23      | 10.98     | 71.7            | English         | Bridgetown         | Prize Raffle             |
| Belarus        | 99       | Stud./Adult   | 27.66      | 10.49     | 51.5            | Russian         | Minsk              | Payment                  |
| Belgium        | 100      | Student       | 21.54      | 4.87      | 81.0            | French          | Louvain-la-Neuve   | Prize Raffle             |
| Brazil         | 101      | Student       | 23.46      | 6.72      | 67.7            | Portuguese      | Aracaju            | Course Credit            |
| Canada         | 347      | Student       | 21.12      | 4.75      | 79.5            | English         | Waterloo           | Course Credit            |
| Chile          | 113      | Student       | 22.39      | 8.61      | 74.1            | Spanish         | Santiago           | Payment                  |
| China          | 247      | Student       | 18.53      | 1.11      | 70.9            | Simpl. Chinese  | Wuhan              | Course Credit            |
| Colombia       | 70       | Student       | 22.54      | 6.00      | 80.0            | Spanish         | Cali               | Payment                  |
| Costa Rica     | 71       | Student       | 24.08      | 3.27      | 66.2            | Spanish         | San Jose           | Payment                  |
| Croatia        | 115      | Student       | 22.56      | 3.15      | 80.9            | Croatian        | Zagreb             | Payment                  |
| Cyprus         | 177      | Student       | 25.30      | 6.40      | 64.8            | Greek           | Nicosia            | Payment                  |
| Czech Rep.     | 101      | Stud./Adult   | 25.28      | 5.15      | 36.6            | Czech           | Various            | Payment                  |
| Denmark        | 173      | Student       | 24.42      | 5.31      | 79.1            | English         | Copenhagen, Aarhus | Course Credit            |
| DRC            | 49       | Student       | 24.33      | 3.55      | 12.2            | French          | Lubumbashi         | Payment                  |
| Egypt          | 1,063    | Student       | 19.78      | 1.93      | 74.8            | Arabic          | Menofia            | Payment                  |
| Estonia        | 99       | Student       | 23.51      | 4.87      | 51.5            | Estonian        | Various            | Payment                  |
| Ethiopia       | 43       | Student       | 21.42      | 3.56      | 55.8            | Amharic, Oromo  | Various            | Payment                  |
| Finland        | 102      | Stud./Adult   | 29.12      | 8.62      | 32.4            | Finnish         | Various            | Payment                  |
| France         | 142      | Stud./Adult   | 25.82      | 9.79      | 76.1            | French          | Occitanie, Paris   | Volun., Raffle           |
| Germany        | 71       | Student       | 21.61      | 4.51      | 84.5            | German          | Mannheim           | Course Credit            |
| Ghana          | 137      | Student       | 20.96      | 3.82      | 69.9            | English         | Accra              | Payment                  |
| Greece         | 202      | Student       | 27.18      | 11.09     | 64.5            | Greek           | Athens             | Volun., Payment          |
| Greenland      | 54       | Student       | 29.83      | 7.05      | 86.5            | Greenlandic     | Nuuk               | Prize Raffle             |
| Guyana         | 197      | Student       | 29.52      | 7.83      | 86.8            | English         | Georgetown         | Prize Raffle             |
| Hong Kong      | 516      | Student       | 22.64      | 5.49      | 77.5            | English         | Hong Kong          | Payment                  |
| Hungary        | 127      | Student       | 26.65      | 9.25      | 82.7            | Hungarian       | Pécs               | Volunteer                |
| Iceland        | 130      | Student       | 25.38      | 7.25      | 56.9            | Icelandic       | Reykjavik          | Prize Raffle             |
| India          | 103      | Student       | 23.94      | 6.90      | 40.6            | English         | Chandigarh         | Payment                  |

<sup>a</sup>CC = Course credit. “Various” = internet samples, where participants lived in many places.

**Table S7. Sample demographics by culture (part 2).**

| <b>Culture</b>  | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b> | <b>Site</b>            | <b>Compensation</b>      |
|-----------------|----------|---------------|------------|-----------|-----------------|-----------------|------------------------|--------------------------|
| Indonesia       | 354      | Student       | 20.49      | 3.20      | 69.7            | Indonesian      | Pekanbaru              | Payment                  |
| Iran            | 78       | Student       | 29.53      | 5.43      | 57.7            | Persian         | Tehran                 | Volunteer                |
| Iraq            | 51       | Student       | 20.86      | 2.63      | 33.3            | Arabic          | Various                | Payment                  |
| Ireland         | 108      | Student       | 19.59      | 2.78      | 79.4            | English         | Limerick               | Volunteer                |
| Israel          | 128      | Student       | 25.63      | 2.93      | 50.8            | Hebrew          | Various                | Payment                  |
| Italy           | 81       | Student       | 29.44      | 10.39     | 79.0            | Italian         | Naples                 | Volunteer                |
| Ivory Coast     | 107      | Student       | 20.34      | 3.14      | 43.0            | French          | Abidjan                | Payment                  |
| Japan           | 173      | Student       | 20.00      | 3.26      | 45.6            | Japanese        | Hokkaido, Tokyo        | Course Credit            |
| Kazakhstan      | 235      | Stud./Adult   | 27.34      | 9.15      | 50.9            | Russian         | Pavlodar               | Course Credit            |
| Kenya           | 333      | Student       | 22.52      | 2.83      | 46.8            | English         | Nairobi, Various       | Payment                  |
| Kuwait          | 58       | Student       | 23.91      | 5.49      | 75.9            | Arabic          | Kuwait                 | Course Credit            |
| Latvia          | 69       | Student       | 21.52      | 4.04      | 46.4            | Latvian         | Various                | Payment                  |
| Lebanon         | 113      | Student       | 18.69      | 1.12      | 54.9            | English         | Beirut                 | Volunteer                |
| Macau           | 133      | Student       | 24.09      | 6.35      | 68.4            | Trad. Chinese   | Macau                  | Payment                  |
| Madagascar      | 100      | Student       | 20.40      | 2.37      | 50.0            | French          | Antananarivo           | Payment                  |
| Malawi          | 65       | Student       | 22.80      | 2.92      | 35.4            | English         | Blantyre               | Payment                  |
| Malaysia        | 106      | Student       | 24.66      | 5.76      | 60.4            | English, Malay  | Kuala Lumpur           | Payment                  |
| Mali            | 92       | Student       | 21.53      | 2.02      | 25.0            | French          | Bamako                 | Payment                  |
| Mexico          | 125      | Student       | 19.75      | 1.39      | 57.5            | Spanish         | Mex. City, Guadalajara | Course Credit            |
| Mongolia        | 194      | Student       | 23.06      | 6.75      | 64.8            | Mongolian       | Ulaanbaatar            | Volunteer                |
| Morocco         | 69       | Student       | 18.57      | 0.72      | 85.5            | French          | Tangier                | Course Credit            |
| Mozambique      | 64       | Student       | 20.61      | 2.29      | 32.8            | Portuguese      | Maputo                 | Payment                  |
| Myanmar         | 113      | Student       | 21.33      | 3.22      | 52.2            | Burmese         | Magway, Various        | Payment                  |
| Namibia         | 127      | Student       | 23.85      | 5.25      | 56.7            | English         | Kavango                | Payment                  |
| Nepal           | 133      | Student       | 29.90      | 10.01     | 62.1            | Nepali          | Kathmandu              | Payment                  |
| Netherlands     | 367      | Student       | 21.22      | 4.20      | 62.9            | Dutch           | Amsterdam, Various     | Payment, CC <sup>a</sup> |
| New Zealand     | 107      | Student       | 21.96      | 4.67      | 51.9            | English         | Wellington             | Payment                  |
| Nicaragua       | 51       | Student       | 20.84      | 2.01      | 49.0            | Spanish         | Managua                | Payment                  |
| Nigeria         | 54       | Student       | 20.07      | 1.91      | 66.7            | English         | Lagos                  | Payment                  |
| North Macedonia | 71       | Stud./Adult   | 31.30      | 8.42      | 70.0            | Macedonian      | Tetovo                 | Prize Raffle             |
| Norway          | 151      | Student       | 23.68      | 5.04      | 76.8            | Norweg., Eng.   | Oslo                   | Course Credit            |
| Oman            | 49       | Student       | 20.15      | 1.80      | 42.9            | Arabic          | Various                | Payment                  |
| Pakistan        | 429      | Student       | 23.61      | 7.07      | 50.8            | English         | Haripur                | Donation to Charity      |
| Peru            | 90       | Student       | 26.81      | 9.39      | 62.4            | Spanish         | Lima                   | Volunteer                |
| Philippines     | 75       | Student       | 25.24      | 7.05      | 86.3            | English         | Cebu                   | Volunteer                |

<sup>a</sup>CC = Course credit. “Various” = internet samples, where participants lived in many places.

**Table S8. Sample demographics by culture (part 3).**

| <b>Culture</b>    | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b>         | <b>Site</b>                                                              | <b>Compensation</b>         |
|-------------------|----------|---------------|------------|-----------|-----------------|-------------------------|--------------------------------------------------------------------------|-----------------------------|
| Poland            | 101      | Student       | 21.85      | 3.96      | 26.7            | Polish                  | Various                                                                  | Payment                     |
| Portugal          | 117      | Student       | 26.32      | 8.26      | 78.6            | Portuguese              | Lisbon                                                                   | Prize Raffle                |
| Puerto Rico       | 97       | Student       | 21.55      | 2.61      | 71.1            | Spanish                 | Various                                                                  | Payment                     |
| Romania           | 164      | Student       | 24.31      | 7.53      | 81.0            | Romanian                | Botosani                                                                 | Payment                     |
| Russia            | 209      | Stud./Adult   | 30.44      | 13.51     | 56.0            | Russian                 | Moscow                                                                   | CC <sup>a</sup> , Volunteer |
| Saudi Arabia      | 170      | Student       | 24.54      | 7.29      | 62.9            | Arabic                  | Tabuk                                                                    | Volunteer                   |
| Senegal           | 400      | Student       | 21.92      | 2.79      | 59.9            | French                  | Dakar                                                                    | Course Credit               |
| Serbia            | 127      | Stud./Adult   | 36.21      | 11.38     | 83.3            | Serbian                 | Various                                                                  | Volunteer                   |
| Singapore         | 67       | Student       | 20.72      | 1.78      | 68.7            | English                 | Singapore                                                                | Course Credit               |
| Slovakia          | 132      | Student       | 22.31      | 2.67      | 65.2            | Slovak                  | Various                                                                  | Payment                     |
| Slovenia          | 101      | Stud./Adult   | 25.87      | 8.58      | 27.7            | Slovenian               | Various                                                                  | Payment                     |
| South Africa      | 160      | Student       | 22.44      | 2.93      | 65.0            | English                 | Johannesburg, Pretoria                                                   | Volunteer                   |
| South Korea       | 100      | Student       | 21.15      | 2.50      | 64.0            | Korean                  | Seoul                                                                    | Course Credit               |
| Spain             | 53       | Student       | 22.53      | 1.51      | 50.9            | Spanish                 | Madrid                                                                   | Volunteer                   |
| Sri Lanka         | 104      | Student       | 23.41      | 1.42      | 77.9            | Sinhala                 | Colombo                                                                  | Payment                     |
| Sweden            | 65       | Student       | 24.78      | 5.86      | 36.9            | Swedish                 | Various                                                                  | Payment                     |
| Switzerland       | 57       | Student       | 25.44      | 5.87      | 40.4            | German, French, Italian | Various                                                                  | Payment                     |
| Taiwan            | 192      | Stud./Adult   | 37.44      | 8.95      | 58.3            | Trad. Chinese           | Taipei                                                                   | Volunteer                   |
| Thailand          | 341      | Student       | 20.45      | 2.34      | 75.1            | Thai                    | Udon Thani, Bangkok                                                      | Payment                     |
| Trinidad & Tobago | 51       | Student       | 21.16      | 2.03      | 82.4            | English                 | San Juan                                                                 | Payment                     |
| Turkey            | 156      | Student       | 21.75      | 1.41      | 53.8            | Turkish                 | Istanbul                                                                 | CC <sup>a</sup> , Volunteer |
| Uganda            | 58       | Student       | 21.93      | 1.66      | 67.2            | English                 | Kampala                                                                  | Payment                     |
| UK                | 538      | Student       | 29.13      | 10.19     | 55.4            | English                 | London, Various                                                          | Payment                     |
| Ukraine           | 228      | Stud./Adult   | 34.00      | 13.74     | 73.5            | Ukrainian               | Kyiv                                                                     | CC <sup>a</sup> , Volunteer |
| United States     | 925      | Student       | 20.39      | 3.52      | 66.7            | English                 | Ann Arbor, Greenville, LA, Richardson, Charlottesville, Chicago, Various | Payment, Course Credit      |
| Uzbekistan        | 99       | Student       | 21.87      | 1.98      | 65.3            | Russian                 | Tashkent                                                                 | Volunteer                   |
| Venezuela         | 90       | Student       | 20.61      | 1.28      | 56.7            | Spanish                 | Various                                                                  | Payment                     |
| Vietnam           | 139      | Student       | 21.86      | 3.61      | 80.6            | Vietnamese              | Ho Chi Minh                                                              | Payment                     |
| Zambia            | 102      | Student       | 24.87      | 6.42      | 46.5            | English                 | Lusaka                                                                   | Payment                     |
| Zimbabwe          | 867      | Student       | 32.01      | 7.66      | 50.2            | English                 | Harare                                                                   | Payment                     |

<sup>a</sup>CC = Course credit. “Various” = internet samples, where participants lived in many places.

**Table S9A. Correlations between responsibilism and previous collectivism scales at the individual level.**

|                        |                                       | <i>r</i>     | <i>P</i> | <i>P</i><br>Adjusted <sup>c</sup> | 95% <i>CI</i> |       | Citations |
|------------------------|---------------------------------------|--------------|----------|-----------------------------------|---------------|-------|-----------|
| Singelis, 1994         | Interdependence                       | 0.04         | 0.647    | 0.703                             | -0.12         | 0.18  | 5,527     |
|                        | Independence                          | 0.08         | 0.316    | 0.429                             | -0.07         | 0.22  |           |
|                        | Inter. - Independence                 | -0.03        | 0.666    | 0.703                             | -0.18         | 0.12  |           |
| Singelis et al., 1995  | Collectivism (Hor. + Vert.)           | <b>0.13</b>  | < 0.001  | < 0.001                           | 0.07          | 0.20  | 3,564     |
|                        | Horizontal Collectivism               | <b>-0.27</b> | < 0.001  | < 0.001                           | -0.34         | -0.21 |           |
|                        | Vertical Collectivism                 | <b>0.41</b>  | < 0.001  | < 0.001                           | 0.35          | 0.47  |           |
|                        | Horizontal Individualism              | 0.04         | 0.218    | 0.345                             | -0.03         | 0.11  |           |
|                        | Vertical Individualism                | <b>0.36</b>  | < 0.001  | < 0.001                           | 0.30          | 0.42  |           |
|                        | Collect. – Individualism <sup>a</sup> | <b>-0.13</b> | < 0.001  | < 0.001                           | -0.20         | -0.06 |           |
| Oyserman, 1993         | Collectivism                          | 0.06         | 0.577    | 0.685                             | -0.12         | 0.20  | 594       |
|                        | Individualism                         | <b>-0.33</b> | 0.001    | 0.003                             | -0.50         | -0.14 |           |
|                        | Coll. – Individualism <sup>b</sup>    | <b>0.28</b>  | 0.004    | 0.010                             | 0.09          | 0.45  |           |
| Cross et al., 2000     | Relational Self-Construct             | <b>-0.29</b> | < 0.001  | < 0.001                           | -0.37         | -0.20 | 1,743     |
| Hui, 1988              | Collectivism                          | -0.13        | 0.181    | 0.313                             | -0.32         | 0.06  | 1,602     |
| Earley, 1993           | Collectivism                          | <b>-0.26</b> | 0.011    | 0.023                             | -0.44         | -0.06 | 1,624     |
| Vignoles et al., 2016  | Interdependence <sup>b</sup>          | 0.10         | 0.312    | 0.429                             | -0.10         | 0.29  | 527       |
| Gudykunst et al., 1996 | Interdependence                       | -0.03        | 0.758    | 0.758                             | -0.23         | 0.17  | 1,936     |
|                        | Independence                          | -0.15        | 0.146    | 0.277                             | -0.34         | 0.05  |           |
|                        | Inter. - Independence                 | 0.09         | 0.397    | 0.503                             | -0.12         | 0.28  |           |

Note: The color shading represents whether the correlation with responsibilism is in the logical direction (such as, one collectivism scale positively correlating with responsibilism).

**Red shading:** Correlates in the opposite direction from expectations

**Green shading:** Correlates in the same direction as expectations

Correlations in **bold** are statistically significant.

[Table S27A](#) shows the sample sizes and sources for each scale. <sup>a</sup>We calculated an overall collectivism score for the Singelis scale by adding the two collectivism scores and subtracting the two individualism scores.

<sup>b</sup>We created an overall interdependence score for the scales of Oyserman and Vignoles, although they do not advocate using sum scores. This table reports all of the facets of Oyserman's scale. [Table S10](#) reports the correlations for each of the facets for Vignoles' scale. Citations are from Google Scholar on March 26, 2024.

<sup>c</sup>We adjusted for multiple comparisons using the Benjamini-Hochberg method.



**Table S9B. Correlations between responsibilism and previous collectivism scales at the culture level.**

|                                    |                                    | <i>r</i>     | <i>P</i> | <i>P</i><br>Adjusted <sup>c</sup> | 95% <i>CI</i> |       | Citations |
|------------------------------------|------------------------------------|--------------|----------|-----------------------------------|---------------|-------|-----------|
| Singelis, 1994                     | Interdependence                    | 0.09         | 0.596    | 0.766                             | -0.25         | 0.41  | 5,527     |
|                                    | Independence                       | 0.05         | 0.756    | 0.820                             | -0.28         | 0.38  |           |
|                                    | Inter. - Independence              | 0.04         | 0.820    | 0.820                             | -0.30         | 0.37  |           |
| Singelis et al., 1995 <sup>b</sup> | Collectivism (Hor. + Vert.)        | <b>0.39</b>  | 0.005    | 0.009                             | 0.13          | 0.60  | 3,564     |
|                                    | Horizontal Collectivism            | <b>-0.40</b> | 0.004    | 0.009                             | -0.61         | -0.14 |           |
|                                    | Vertical Collectivism <sup>a</sup> | <b>0.71</b>  | < 0.001  | < 0.001                           | 0.53          | 0.82  |           |
| Cross et al., 2000                 | Relational Self-Construal          | <b>-0.35</b> | 0.012    | 0.018                             | -0.57         | -0.08 | 1,743     |
| Earley, 1993                       | Collectivism                       | <b>-0.40</b> | 0.004    | 0.009                             | -0.61         | -0.13 | 1,624     |
| Gudykunst et al., 1996             | Interdependence                    | <b>-0.67</b> | < 0.001  | < 0.001                           | -0.80         | -0.49 | 1,936     |

Note: The color shading represents whether the correlation with responsibilism is in the logical direction (such as, one collectivism scale positively correlating with responsibilism).

**Red shading:** Correlates in the opposite direction from expectations

**Green shading:** Correlates in the same direction as expectations

Correlations in **bold** are statistically significant.

[Table S27A](#) shows the sample sizes and sources for each scale. The Singelis 1994 scores come from a recently published study<sup>33</sup>. <sup>a</sup>From our perspective, the Singelis 1995 collectivism scale has the items that come closest to responsibilism. That is because six items mention specific relationships (mostly parents and children). Several of the items mention responsibilities and social roles, with concepts like taking care of elderly parents, “duty,” and “sacrifice.” Because the scale contained items that (a) specified relationships and (b) described roles and responsibilities in those relationships, we used three items from this scale in the responsibilism scale. However, the scale has the downside of using the vague word “group,” which we argue does not tap into responsibilism. This is because groups often include weak ties and strangers, such as with ethnic groups, companies, and fans of the same sports team. Despite its strengths, the Singelis scale also has some of the items that fit most closely with the misleading warm fuzzy version of collectivism, such as, “I feel good when I cooperate with others” and “To me, pleasure is spending time with others.” The fact that the Singelis 1995 scale mixes duties in close relationships (in its vertical collectivism sub-scale) and warm fuzzies in vague relationships (in its horizontal collectivism sub-scale) while calling both “collectivism” is a fitting example of people’s conflicting intuitions about collectivism. <sup>b</sup>The individual-level data in Table S9A includes the individualism sub-scales for this scale. However, for our data collection in 50 cultures, we did not have the space capacity to include the Singelis independence sub-scales. Citations are from Google Scholar on March 26, 2024. <sup>c</sup>This column reports *P* values adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S9C. Correlations between responsibilism and previous collectivism scales at the culture level after controlling for agreeableness.**

|                                    |                                    | <i>r</i>     | <i>P</i> | <i>P</i><br>Adjusted <sup>c</sup> | 95% <i>CI</i> |       | Citations |
|------------------------------------|------------------------------------|--------------|----------|-----------------------------------|---------------|-------|-----------|
| Singelis, 1994                     | Interdependence                    | 0.15         | 0.452    | 0.509                             | -0.25         | 0.51  | 5,527     |
|                                    | Independence                       | 0.25         | 0.219    | 0.282                             | -0.15         | 0.58  |           |
|                                    | Inter. - Independence              | 0.002        | 0.991    | 0.991                             | -0.39         | 0.39  |           |
| Singelis et al., 1995 <sup>b</sup> | Collectivism (Hor. + Vert.)        | 0.38         | 0.041    | 0.092                             | 0.02          | 0.65  | 3,564     |
|                                    | Horizontal Collectivism            | -0.34        | 0.063    | 0.094                             | -0.63         | 0.02  |           |
|                                    | Vertical Collectivism <sup>a</sup> | <b>0.59</b>  | 0.001    | 0.009                             | 0.29          | 0.78  |           |
| Cross et al., 2000                 | Relational Self-Construct          | -0.39        | 0.032    | 0.092                             | -0.66         | -0.04 | 1,743     |
| Earley, 1993                       | Collectivism                       | -0.35        | 0.060    | 0.094                             | -0.63         | 0.02  | 1,624     |
| Gudykunst et al., 1996             | Interdependence                    | <b>-0.55</b> | 0.002    | 0.009                             | -0.76         | -0.24 | 1,936     |

Note: This table tests culture-level correlations using agreeableness data across cultures from a previous study<sup>111</sup>. The sample of countries differs between our study and the personality study, which means there are fewer overlapping countries. The sample size here is 27 (except for the Singelis 1994 scale, which has 23 countries). The color shading represents whether the correlation with responsibilism is in the logical direction (such as, one collectivism scale positively correlating with responsibilism).

**Red shading:** Correlates in the opposite direction from expectations

**Green shading:** Correlates in the same direction as expectations

Correlations in **bold** are statistically significant.

[Table S27A](#) shows the sample sizes and sources for each scale. The Singelis 1994 scores come from a recently published study<sup>33</sup>. <sup>a</sup>From our perspective, the Singelis 1995 collectivism scale has the items that come closest to responsibilism. That is because six items mention specific relationships (mostly parents and children). Several of the items mention responsibilities and social roles, with concepts like taking care of elderly parents, “duty,” and “sacrifice.” Because the scale contained items that (a) specified relationships and (b) described roles and responsibilities in those relationships, we used three items from this scale in the responsibilism scale. However, the scale has the downside of using the vague word “group,” which we argue does not tap into responsibilism. This is because groups often include weak ties and strangers, such as with ethnic groups, companies, and fans of the same sports team. Despite its strengths, the Singelis scale also has some of the items that fit most closely with the misleading warm fuzzy version of collectivism, such as, “I feel good when I cooperate with others” and “To me, pleasure is spending time with others.” The fact that the Singelis 1995 scale mixes duties in close relationships (in its vertical collectivism sub-scale) and warm fuzzies in vague relationships (in its horizontal collectivism sub-scale) while calling both “collectivism” is a fitting example of people’s conflicting intuitions about collectivism. <sup>b</sup>The individual-level data in Table S9A includes the individualism sub-scales for this scale. However, for our data collection in 50 cultures, we did not have the space capacity to include the Singelis independence sub-scales. <sup>c</sup>This column reports *P* values adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S10. Divergent validity tests show responsibilism is different from Vignoles' self-construal scale components.**

| Correlations of responsibilism with... | <i>r</i> | <i>P</i> | <i>P</i><br>Adjusted <sup>a</sup> | <i>95% CI</i> |      |
|----------------------------------------|----------|----------|-----------------------------------|---------------|------|
| Difference                             | 0.08     | 0.426    | 0.784                             | -0.12         | 0.27 |
| Self-Containment                       | -0.06    | 0.537    | 0.784                             | -0.26         | 0.14 |
| Self-Direction                         | -0.14    | 0.157    | 0.707                             | -0.33         | 0.06 |
| Self-Reliance                          | -0.06    | 0.556    | 0.784                             | -0.25         | 0.14 |
| Self-Expression                        | -0.05    | 0.610    | 0.784                             | -0.24         | 0.15 |
| Self-Interest                          | 0.03     | 0.780    | 0.878                             | -0.17         | 0.22 |
| Consistency                            | -0.20    | 0.046    | 0.414                             | -0.38         | 0.00 |
| De-Contextualized Self                 | 0.002    | 0.985    | 0.985                             | -0.19         | 0.20 |
| Sum Interdependence Vignoles           | 0.10     | 0.312    | 0.784                             | -0.10         | 0.29 |

*N* = 101

Note: This table tests whether the responsibilism scale is different from the sub-facets of Vignoles' self-construal scale<sup>45</sup>. Participants were 101 college students in the UK on the platform Prolific. The final row presents the sum score of all the sub-facets of scale. We present this because researchers have used the scale this way. However, Vignoles does not use these sum scores. Instead, he keeps the sub-facets separate. Respondents took an abbreviated version of the responsibilism scale, so it does not include all items of the scale. <sup>a</sup>The *P* values in this column are adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S11A. Correlations between responsibility and personality.**

|                   | <i>r</i>     | <i>P</i> | <i>P</i><br>Adjusted <sup>a</sup> | 95% CI |       |
|-------------------|--------------|----------|-----------------------------------|--------|-------|
| Extraversion      | -0.09        | 0.262    | 0.328                             | -0.23  | 0.06  |
| Neuroticism       | -0.12        | 0.127    | 0.212                             | -0.26  | 0.03  |
| Agreeableness     | <b>-0.22</b> | 0.003    | 0.015                             | -0.36  | -0.08 |
| Conscientiousness | 0.04         | 0.580    | 0.580                             | -0.11  | 0.19  |
| Openness          | <b>-0.18</b> | 0.017    | 0.043                             | -0.32  | -0.03 |

Note: Participants took the Ten Item Personality Scale used to measure the Big Five personality factors. Participants were 173 students from the participant pool at the University of Texas Dallas. Correlations in **bold** are statistically significant. <sup>a</sup>We adjusted for multiple comparisons using the Benjamini-Hochberg method to reduce the false discovery rate.

**Table S11B. Agreeableness correlated negatively with responsibilism but positively with Singelis's collectivism scores at the individual level.**

|                                                                                               | <i>r</i> | <i>P</i> | <i>P</i><br>Adjusted <sup>a</sup> | 95% CI |       |
|-----------------------------------------------------------------------------------------------|----------|----------|-----------------------------------|--------|-------|
| Correlations of agreeableness with...                                                         |          |          |                                   |        |       |
| Responsibilism                                                                                | -0.22    | 0.003    | 0.012                             | -0.36  | -0.08 |
| Independence (Singelis)                                                                       | -0.04    | 0.589    | 0.589                             | -0.19  | 0.11  |
| Interdependence (Singelis)                                                                    | 0.17     | 0.023    | 0.046                             | 0.03   | 0.31  |
| Inter. – Independence (Singelis)                                                              | 0.15     | 0.045    | 0.060                             | 0.003  | 0.30  |
| Correlations of responsibilism with the Singelis scale<br>after controlling for agreeableness |          |          |                                   |        |       |
| Independence                                                                                  | 0.07     | 0.377    | 0.566                             | -0.08  | 0.22  |
| Interdependence                                                                               | 0.08     | 0.329    | 0.566                             | -0.08  | 0.22  |
| Inter. - Independence                                                                         | 0.001    | 0.991    | 0.991                             | -0.15  | 0.15  |

Note: This table tests the idea that previous collectivism scales were confounded with agreeableness. The logic is that, if they were confounded, controlling for agreeableness might lead to the expected correlations between previous collectivism scales and responsibilism. However, the correlations were not significant. We were able to test this in a sub-sample of participants that took the responsibilism scale, the Singelis 1994 independence-interdependence scale<sup>2</sup>, and the Ten Item Personality Inventory to measure the Big Five personality factors<sup>112</sup>. Participants were 173 students from the participant pool at the University of Texas Dallas. Correlations in **bold** are statistically significant. [Table S9C](#) analyzes agreeableness at the culture level. <sup>a</sup>This column reports *P* values adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S12A. Previous collectivism scales show contradictory correlations with social status.**

|                        |                              | <i>r</i>    | <i>P</i> | <i>P</i><br>Adjusted <sup>b</sup> | 95% CI |       |
|------------------------|------------------------------|-------------|----------|-----------------------------------|--------|-------|
| Singelis, 1994         | Interdependence              | -0.09       | 0.260    | 0.449                             | -0.23  | 0.06  |
|                        | Independence                 | -0.04       | 0.629    | 0.908                             | -0.19  | 0.11  |
|                        | Inter. - Ind.                | -0.03       | 0.669    | 0.908                             | -0.18  | 0.12  |
| Singelis et al., 1995  | Collectivism (Hor. + Vert.)  | 0.03        | 0.063    | 0.171                             | -0.002 | 0.07  |
|                        | Horizontal Collectivism      | <b>0.05</b> | 0.003    | 0.048                             | 0.02   | 0.09  |
|                        | Vertical Collectivism        | 0.001       | 0.952    | 0.952                             | -0.03  | 0.04  |
|                        | Horizontal Individualism     | -0.04       | 0.223    | 0.424                             | -0.11  | 0.03  |
|                        | Vertical Individualism       | -0.09       | 0.009    | 0.057                             | -0.15  | -0.02 |
|                        | Coll. – Individualism        | 0.03        | 0.412    | 0.652                             | -0.04  | 0.09  |
| Oyserman, 1993         | Collectivism                 | -0.03       | 0.766    | 0.952                             | -0.22  | 0.16  |
|                        | Individualism                | 0.23        | 0.017    | 0.081                             | 0.04   | 0.41  |
|                        | Coll. – Individualism        | -0.19       | 0.053    | 0.168                             | -0.37  | 0.00  |
| Cross, 2000            | Relational Self-Construal    | 0.03        | 0.122    | 0.290                             | -0.01  | 0.06  |
| Hui, 1988              | Collectivism                 | 0.14        | 0.165    | 0.348                             | -0.06  | 0.33  |
| Earley, 1993           | Collectivism                 | 0.04        | 0.024    | 0.091                             | 0.01   | 0.07  |
| Gudykunst et al., 1996 | Independence                 | 0.01        | 0.940    | 0.952                             | -0.19  | 0.21  |
|                        | Interdependence              | <b>0.05</b> | 0.005    | 0.048                             | 0.02   | 0.08  |
|                        | Inter. - Independence        | 0.01        | 0.894    | 0.952                             | -0.19  | 0.21  |
| Vignoles et al., 2016  | Interdependence <sup>a</sup> | -0.02       | 0.812    | 0.952                             | -0.20  | 0.16  |

Note: The color shading represents whether the correlation with social status is in the theoretically predicted direction (higher social status = less collectivistic)<sup>41,42</sup>.

**Red shading:** Correlates in the opposite direction from expectations

**Green shading:** Correlates in the same direction as expectations

Correlations in **bold** are statistically significant.

Participants rated their social status compared to people in their country on a ladder from 0 (*bottom*) to 10 (*top*).

[Tables S27A-E](#) show the sample sizes and sources for each scale. The sample sizes are larger for the four prior collectivism scales that we tested across cultures: the Singelis et al., 1995 collectivism sub-scales; the Gudykunst interdependence sub-scale; the Cross scale; and the Earley scale. <sup>a</sup>We created an overall interdependence score for Vignoles and colleagues' scale. Although researchers have used sum scores this scale, Vignoles does not advocate summing the sub-scales of this scale. <sup>b</sup>We adjusted for multiple comparisons using the Benjamini-Hochberg method to reduce the false discovery rate.

**Table S12B. Previous collectivism scales show contradictory correlations with political conservatism.**

|                        |                              | <i>r</i>     | <i>P</i> | <i>P</i><br>Adjusted <sup>b</sup> | 95% CI |       |
|------------------------|------------------------------|--------------|----------|-----------------------------------|--------|-------|
| Singelis, 1994         | Interdependence              | 0.17         | 0.032    | 0.068                             | 0.02   | 0.31  |
|                        | Independence                 | 0.14         | 0.077    | 0.133                             | -0.02  | 0.28  |
|                        | Inter. - Ind.                | 0.01         | 0.861    | 0.976                             | -0.14  | 0.17  |
| Singelis et al., 1995  | Collectivism (Hor. + Vert.)  | 0.03         | 0.063    | 0.120                             | -0.002 | 0.07  |
|                        | Horizontal Collectivism      | <b>-0.09</b> | < 0.001  | < 0.001                           | -0.13  | -0.06 |
|                        | Vertical Collectivism        | <b>0.15</b>  | < 0.001  | < 0.001                           | 0.11   | 0.18  |
|                        | Horizontal Individualism     | -0.08        | 0.023    | 0.062                             | -0.14  | -0.01 |
|                        | Vertical Individualism       | 0.08         | 0.027    | 0.064                             | 0.01   | 0.14  |
|                        | Collect. – Individualism     | 0.08         | 0.020    | 0.062                             | 0.01   | 0.14  |
| Oyserman, 1993         | Collectivism                 | 0.01         | 0.954    | 0.978                             | -0.19  | 0.20  |
|                        | Individualism                | -0.15        | 0.138    | 0.219                             | -0.33  | 0.05  |
|                        | Coll. - Ind.                 | 0.11         | 0.266    | 0.361                             | -0.08  | 0.29  |
| Cross, 2000            | Relational Self-Construal    | <b>-0.13</b> | < 0.001  | < 0.001                           | -0.16  | -0.10 |
| Hui, 1988              | Collectivism                 | -0.02        | 0.873    | 0.976                             | -0.21  | 0.18  |
| Earley, 1993           | Collectivism                 | <b>-0.07</b> | < 0.001  | < 0.001                           | -0.10  | -0.03 |
| Gudykunst et al., 1996 | Independence                 | -0.003       | 0.978    | 0.978                             | -0.20  | 0.20  |
|                        | Interdependence              | <b>-0.14</b> | < 0.001  | < 0.001                           | -0.17  | -0.11 |
|                        | Inter. - Ind.                | 0.14         | 0.179    | 0.262                             | -0.06  | 0.33  |
| Vignoles et al., 2016  | Interdependence <sup>a</sup> | 0.04         | 0.703    | 0.890                             | -0.14  | 0.21  |

Note: The color shading represents whether the correlation with social conservatism is in the same direction as the findings of cultural thought style<sup>44,56</sup> and responsibilism (social conservatives are more collectivistic).

**Red shading:** Correlates in the opposite direction from expectations

**Green shading:** Correlates in the same direction as expectations

Correlations in **bold** are statistically significant.

Participants reported rated their social politics from 1 (*very liberal*) to 7 (*very conservative*). We focused on social politics because previous studies have found that social politics are a better predictor of cultural differences than economic politics or overall politics<sup>44,56</sup>. [Tables S27A-E](#) show the sample sizes and sources for each scale. The sample sizes are larger for the four prior collectivism scales that we tested across cultures: the Singelis et al., 1995 collectivism sub-scales; the Gudykunst interdependence sub-scale; the Cross scale; and the Earley scale. <sup>a</sup>We created an overall interdependence score for Vignoles and colleagues' scale. Although researchers have used sum scores for this scale, Vignoles does not advocate summing the sub-scales of this scale. <sup>b</sup>This column reports *P* values adjusted for multiple comparisons using the Benjamini-Hochberg method.



**Table S13A. Survey responses from the US and China used for comparison with researchers' guesses (part 1).**

| Red shading = sig. diff.<br>opposing warm fuzzy,<br>green = favoring warm<br>fuzzy                         |       | <i>Avg.</i> | <i>SEM</i> | <i>t</i> | <i>P</i> | <i>P</i><br>Adjusted <sup>a</sup> | <i>d</i><br>[95% CI] | Scale                                    | Based on Scale,<br>Collectivists<br>Should<br>Agree... | Taps into<br>Warm Fuzzy<br>Intuitions? |
|------------------------------------------------------------------------------------------------------------|-------|-------------|------------|----------|----------|-----------------------------------|----------------------|------------------------------------------|--------------------------------------------------------|----------------------------------------|
| I often have the feeling that my relationships with others are more important than my own accomplishments. | US    | 4.56        | 0.13       | 4.08     | < 0.001  | < 0.001                           | 0.59                 | Singelis, 1994                           | More                                                   | Yes                                    |
|                                                                                                            | China | 3.67        | 0.17       |          |          |                                   | [0.30, 0.88]         |                                          |                                                        |                                        |
| It is important that I do my job better than others.                                                       | US    | 4.96        | 0.11       | 0.13     | 0.895    | 0.895                             | 0.02                 | Singelis et al., 1995                    | Less                                                   | Yes                                    |
|                                                                                                            | China | 4.93        | 0.13       |          |          |                                   | [-0.27, 0.31]        |                                          |                                                        |                                        |
| Winning is everything.                                                                                     | US    | 2.30        | 0.13       | -6.06    | < 0.001  | < 0.001                           | -0.89                | Singelis et al., 1995                    | Less                                                   | Yes                                    |
|                                                                                                            | China | 3.53        | 0.16       |          |          |                                   | [-1.19, -0.59]       |                                          |                                                        |                                        |
| We should keep our aging parents with us at home.                                                          | US    | 4.12        | 0.13       | -4.47    | < 0.001  | < 0.001                           | -0.66                | Singelis et al., 1995;<br>Responsibilism | More                                                   | No                                     |
|                                                                                                            | China | 5.07        | 0.16       |          |          |                                   | [-0.95, -0.36]       |                                          |                                                        |                                        |
| The well-being of my coworkers is important to me.                                                         | US    | 5.45        | 0.09       | 8.10     | < 0.001  | < 0.001                           | 1.32                 | Singelis et al., 1995                    | More                                                   | Yes                                    |
|                                                                                                            | China | 3.88        | 0.17       |          |          |                                   | [1.00, 1.63]         |                                          |                                                        |                                        |
| To me, pleasure is spending time with others.                                                              | US    | 5.11        | 0.12       | 2.93     | 0.004    | 0.007                             | 0.44                 | Singelis et al., 1995                    | More                                                   | Yes                                    |
|                                                                                                            | China | 4.53        | 0.16       |          |          |                                   | [0.15, 0.74]         |                                          |                                                        |                                        |
| I feel good when I cooperate with others.                                                                  | US    | 5.82        | 0.09       | 5.16     | < 0.001  | < 0.001                           | 0.76                 | Singelis et al., 1995                    | More                                                   | Yes                                    |
|                                                                                                            | China | 5.09        | 0.11       |          |          |                                   | [0.46, 1.06]         |                                          |                                                        |                                        |
| It often happens that the interests of my group coincide with my own interests.                            | US    | 4.94        | 0.12       | 1.76     | 0.080    | 0.095                             | 0.26                 | Oyserman, 1993                           | More                                                   | Yes                                    |
|                                                                                                            | China | 4.59        | 0.15       |          |          |                                   | [-0.03, 0.54]        |                                          |                                                        |                                        |
| I often do "my own thing."                                                                                 | US    | 5.90        | 0.09       | 1.75     | 0.082    | 0.095                             | 0.26                 | Singelis et al., 1995                    | Less                                                   | Yes                                    |
|                                                                                                            | China | 5.64        | 0.11       |          |          |                                   | [-0.03, 0.55]        |                                          |                                                        |                                        |

Note: Participants were 119 college students in the US and 76 in China. Sample sizes differ slightly across scales because we randomly assigned participants to scales. "Taps into warm fuzzy intuitions" means that this item (a) does not specify the relationship and instead uses vague terms like "others" or "group" or (b) is based on the idea that collectivistic cultures are nice to or feel warm feelings about people in general, whereas individualistic cultures are competitive and self-focused. <sup>a</sup>This column reports *P* values adjusted for multiple comparisons using the Benjamini-Hochberg method. The items in rows 5 and 6 showed unequal variance across cultures, so we adjusted the degrees of freedom. Supplemental Section 38 explains this in more detail.

**Table S13B. Survey responses from the US and China used for comparison with researchers' guesses (part 2).**

| Red shading = sig. diff opposing warm fuzzy,<br>green = favoring warm fuzzy                                                                                                                                                                                                                                                |       |      | Avg. | SEM   | <i>t</i> | <i>P</i> | <i>P</i><br>Adjusted <sup>a</sup> | <i>d</i><br>[95% CI] | Scale                 | Based on Scale,<br>Collectivists<br>Should Agree... | Taps into<br>Warm Fuzzy<br>Intuitions? |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|------|------|-------|----------|----------|-----------------------------------|----------------------|-----------------------|-----------------------------------------------------|----------------------------------------|
|                                                                                                                                                                                                                                                                                                                            |       |      |      |       |          |          |                                   |                      |                       |                                                     |                                        |
| I have subdued my personal opinions in order to be part of my group.                                                                                                                                                                                                                                                       | US    | 4.06 | 0.15 | 2.14  | 0.034    | 0.046    |                                   | 0.29                 | Oyserman, 1993        | More                                                | Yes                                    |
|                                                                                                                                                                                                                                                                                                                            | China | 3.59 | 0.16 |       |          |          |                                   | [0.01, 0.58]         |                       |                                                     |                                        |
| I enjoy being unique and different from others in many ways.                                                                                                                                                                                                                                                               | US    | 5.44 | 0.10 | 2.91  | 0.004    | 0.007    |                                   | 0.43                 | Singelis et al., 1995 | Less                                                | Yes                                    |
|                                                                                                                                                                                                                                                                                                                            | China | 4.92 | 0.15 |       |          |          |                                   | [0.14, 0.72]         |                       |                                                     |                                        |
| One should live one's life independently of others.                                                                                                                                                                                                                                                                        | US    | 4.62 | 0.14 | -4.37 | < 0.001  | < 0.001  |                                   | -0.60                | Singelis et al., 1995 | Less                                                | Yes                                    |
|                                                                                                                                                                                                                                                                                                                            | China | 5.45 | 0.12 |       |          |          |                                   | [-0.89, -0.30]       |                       |                                                     |                                        |
| You are working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You've been working on a sales account that could be a big accomplishment. The co-worker asks about the account. I would feel comfortable telling this co-worker about the account. | US    | 4.13 | 0.16 | 2.61  | 0.010    | 0.015    |                                   | 0.37                 | Responsibilism        | Less*                                               | Yes                                    |
|                                                                                                                                                                                                                                                                                                                            | China | 3.51 | 0.17 |       |          |          |                                   | [0.08, 0.66]         |                       |                                                     |                                        |
| If my parents have a nice little thing that happened to them today (for example, they saw a very pretty flower or their co-worker complimented them), my parents would tell me the little story.                                                                                                                           | US    | 5.12 | 0.16 | 0.67  | 0.507    | 0.543    |                                   | 0.10                 | Responsibilism        | Less*                                               | Yes                                    |
|                                                                                                                                                                                                                                                                                                                            | China | 4.96 | 0.17 |       |          |          |                                   | [-0.19, 0.39]        |                       |                                                     |                                        |
| Imagine you live in an apartment with four roommates that you didn't know before you moved in. One day, two friends of your roommate come over. Your roommate says, "My room isn't big enough for both of my friends. Can one sleep in your room?" I would happily agree.                                                  | US    | 2.49 | 0.15 | -2.96 | 0.003    | 0.006    |                                   | -0.43                | Responsibilism        | Less*                                               | Yes                                    |
|                                                                                                                                                                                                                                                                                                                            | China | 3.20 | 0.19 |       |          |          |                                   | [-0.73, -0.14]       |                       |                                                     |                                        |

Note: \*In collectivistic cultures, agreement should be lower based on responsibilistic intuitions but higher based on warm fuzzy intuitions about collectivism.

<sup>a</sup>This column reports *P* values that are adjusted for multiple comparisons using the Benjamini-Hochberg method. The items in rows 1, 3, and 5 showed unequal variance across cultures, so we adjusted the degrees of freedom. Supplemental Section 38 explains this in more detail.

**Table S14. Demographics for the sample of researchers for the guessing survey.**

|                                            |                    | % of Sample | <i>N</i> | Correct Guesses |
|--------------------------------------------|--------------------|-------------|----------|-----------------|
| Grew Up in...                              | East Asia          | 17.5%       | 7        | 20.4%           |
|                                            | West               | 60.0%       | 24       | 23.5%           |
|                                            | Other              | 22.5%       | 9        | 20.6%           |
| Lives in...                                | East Asia          | 12.5%       | 5        | 25.7%           |
|                                            | West               | 72.5%       | 29       | 21.9%           |
|                                            | Other              | 15.0%       | 6        | 21.4%           |
| Career Position                            | Graduate Student   | 5.0%        | 2        | 21.4%           |
|                                            | Post-Doc           | 17.5%       | 7        | 17.4%           |
|                                            | Asst./Assoc. Prof. | 35.0%       | 14       | 21.9%           |
|                                            | Full Professor     | 35.0%       | 14       | 24.5%           |
|                                            | Other              | 7.5%        | 3        | 26.2%           |
| Self-Identifies as a Cultural Psychologist |                    | 35.0%       | 14       | 21.9%           |
| Has Taught Cultural Psychology             |                    | 42.5%       | 17       | 22.7%           |
| <b>Total</b>                               |                    |             | 40       | 22.3%           |

Note: We calculated correct guesses about warm fuzzy survey items based on whether the differences between the US and China were statistically significant in the direction that researchers predicted. [Table S13](#) shows the survey questions and the actual results across cultures. [Fig. S11](#) shows researchers' guesses. If researchers guessed that the two cultures would be "about equal," we counted this as correct if the two cultures did not show a significant difference on the question. Because there were three response options for each question, random guessing would lead to 33% correct guesses on average. The average percent correct was below random guessing. This fits with our idea that people often expect that people in collectivistic cultures will agree with warm fuzzy items. It is logical to hypothesize that accuracy would be higher among people who are more experienced researchers or who grew up in East Asia or other non-Western cultures. However, accuracy differences were small (< 10%) based on where people grew up and their career stage.

**Table S15. Testing whether the five-factor model applies to both individuals and cultures.**

| Model | Description                                | $\chi^2$ (DF)                   | CFI   | RMSEA | SRMR <sub>w</sub> | SRMR <sub>b</sub> | AIC       | BIC       | $\Delta$ CFI | $\Delta$ RMSEA | $\Delta$ SRMR <sub>w</sub> | $\Delta$ SRMR <sub>b</sub> |
|-------|--------------------------------------------|---------------------------------|-------|-------|-------------------|-------------------|-----------|-----------|--------------|----------------|----------------------------|----------------------------|
| 1     | W: Factor structure<br>B: Saturated        | 284.664<br>(48) <sup>***</sup>  | 0.957 | 0.020 | 0.019             | 0.002             | 662000.99 | 663305.69 |              |                |                            |                            |
| 2     | W: Saturated<br>B: Factor structure        | 105.192<br>(48) <sup>***</sup>  | 0.990 | 0.010 | 0.000             | 0.099             | 661689.76 | 662994.46 | 0.033        | -0.010         | -0.019                     | 0.097                      |
| 3     | W: Factor structure<br>B: Factor structure | 419.610<br>(96) <sup>***</sup>  | 0.941 | 0.017 | 0.019             | 0.099             | 662018.83 | 662967.70 | -0.049       | 0.007          | 0.019                      | 0.000                      |
| 4     | M3 with cross-level<br>invariance          | 454.544<br>(110) <sup>***</sup> | 0.937 | 0.016 | 0.020             | 0.100             | 662020.44 | 662865.53 | -0.004       | -0.001         | 0.001                      | 0.001                      |

Note: The final model tests whether the same five-factor structure for the responsibilism scale holds both for individuals and groups. The good fit metrics for this model suggest that the same factor structure fits both for individuals (suggesting we can compare individuals) and groups (suggesting we can compare cultures). The final model surpasses recommended cutoffs for CFI (> 0.90), RMSEA (< 0.08), and SRMR (< 0.08), except that the SRMR for the between-groups model is higher than the cutoff for SRMR (0.10 vs. 0.08).

**W** = Within groups (individual level)

**B** = Between groups (culture level)

**Saturated:** The saturated models test without our proposed factor structure. For example, Model 1 tests our proposed five-factor model at the individual level, but not at the group level.

**CFI** = Comparative fit index. Higher values indicate better model fit. Some researchers suggested a cutoff of > 0.90<sup>108,109</sup>.

**RMSEA** = Root mean squared error of approximation. Better models have less error. Some researchers suggested a cutoff of < 0.08<sup>107</sup>.

**SRMR** = Standardized root mean squared residual. Better models have less error. Some researchers suggested a cutoff of < 0.08<sup>107</sup>.

$\Delta$  = Change in fit indexes compared to the model above.

**AIC** = Akaike information criterion. Lower values indicate better model fit.

**BIC** = Bayesian information criterion. Lower values indicate better model fit.

<sup>\*\*\*</sup> $P < 0.001$

**Table S16. Testing whether the one-factor model applies to both individuals and cultures.**

| Model                                      | $\chi^2$ (DF)                   | CFI   | RMSEA | SRMR <sub>w</sub> | SRMR <sub>b</sub> | AIC       | BIC       | $\Delta$ CFI | $\Delta$ RMSEA | $\Delta$ SRMR <sub>w</sub> | $\Delta$ SRMR <sub>b</sub> |
|--------------------------------------------|---------------------------------|-------|-------|-------------------|-------------------|-----------|-----------|--------------|----------------|----------------------------|----------------------------|
| W: Factor structure<br>B: Saturated        | 351.458<br>(58) <sup>***</sup>  | 0.947 | 0.020 | 0.020             | 0.002             | 662055.55 | 663286.12 |              |                |                            |                            |
| W: Saturated<br>B: Factor structure        | 120.074<br>(58) <sup>***</sup>  | 0.989 | 0.009 | 0.000             | 0.105             | 661675.86 | 662906.42 | 0.042        | -0.011         | -0.020                     | 0.103                      |
| W: Factor structure<br>B: Factor structure | 513.172<br>(116) <sup>***</sup> | 0.928 | 0.017 | 0.020             | 0.106             | 662059.07 | 662859.68 | -0.061       | 0.008          | 0.020                      | 0.001                      |
| M3 with cross-level<br>invariance          | 593.587<br>(130) <sup>***</sup> | 0.916 | 0.017 | 0.020             | 0.317             | 662129.22 | 662826.04 | -0.012       | 0.000          | 0.000                      | 0.211                      |

Note: The final model tests whether the same one-factor structure for the responsibility scale holds both for individuals and groups. The good fit metrics for this model suggest that the same factor structure fits both for individuals (suggesting we can compare individuals) and groups (suggesting we can compare cultures). The final model surpasses recommended cutoffs for CFI ( $> 0.90$ ), RMSEA ( $< 0.08$ ), and SRMR ( $< 0.08$ ), except that the SRMR for the between-groups model is higher than the cutoff for SRMR (0.32 vs. 0.08).

**W** = Within groups (individual level)

**B** = Between groups (culture level)

**Saturated:** The saturated models test without our proposed factor structure. For example, Model 1 tests our proposed five-factor model at the individual level, but not at the group level.

**CFI** = Comparative fit index. Higher values indicate better model fit. Some researchers suggested a cutoff of  $> 0.90$ <sup>108,109</sup>.

**RMSEA** = Root mean squared error of approximation. Better models have less error. Some researchers suggested a cutoff of  $< 0.08$ <sup>107</sup>.

**SRMR** = Standardized root mean squared residual. Better models have less error. Some researchers suggested a cutoff of  $< 0.08$ <sup>107</sup>.

$\Delta$  = Change in fit indexes compared to the model above.

**AIC** = Akaike information criterion. Lower values indicate better model fit.

**BIC** = Bayesian information criterion. Lower values indicate better model fit.

<sup>\*\*\*</sup>  $P < 0.001$

**Table S17. Testing model fit without and with an element representing response style.**

| Model                                    | $\chi^2(df)$    | CFI   | CFI<br>Cutoff | RMSEA | RMSEA<br>Cutoff | SRMR  | SRMR<br>Cutoff | AIC        | BIC        |
|------------------------------------------|-----------------|-------|---------------|-------|-----------------|-------|----------------|------------|------------|
| Five-Factor Model                        | 609.141 (53)*** | 0.942 | > 0.900       | 0.029 | < 0.080         | 0.021 | < 0.080        | 679208.975 | 679698.235 |
| Five-Factor Model<br>with Response Style | 241.482 (39)*** | 0.979 | > 0.900       | 0.021 | < 0.080         | 0.012 | < 0.080        | 678876.674 | 679469.717 |

Note: The bottom model adds a bifactor common-method factor (CMF) to model response style differences. This helps test whether response style makes a large difference in the results. We also used the results to test whether calculating cultures scores using this method of taking into account response style led to better external validity than the simpler method of taking into account response style. The results in [Supplementary Section 18](#) showed that the simpler method of taking into account response style produced culture scores with higher correlations with external markers of collectivism (better external validity).

**CFI** = Comparative fit index.

**RMSEA** = Root mean squared error of approximation.

**SRMR** = Standardized root mean squared residual.

**SRMR** = Standardized root mean squared residual.

**AIC** = Akaike information criterion. Lower values indicate better model fit.

**BIC** = Bayesian information criterion. Lower values indicate better model fit.

\*\*\* $P < 0.001$

**Table S18A. Fit indices for measurement invariance across Western, “sometimes Western,” and non-Western cultures.**

| Model      | $\chi^2$ (DF)                 | CFI   | RMSEA | SRMR  | $\Delta$ CFI | $\Delta$ CFI<br>Cutoff | $\Delta$ RMSEA | $\Delta$ RMSEA<br>Cutoff | $\Delta$ SRMR | $\Delta$ SRMR<br>Cutoff |
|------------|-------------------------------|-------|-------|-------|--------------|------------------------|----------------|--------------------------|---------------|-------------------------|
| Configural | 796.177 (159) <sup>***</sup>  | 0.931 | 0.031 | 0.022 |              |                        |                |                          |               |                         |
| Metric     | 1030.887 (203) <sup>***</sup> | 0.911 | 0.032 | 0.027 | 0.020        | < 0.02                 | 0.001          | < 0.015                  | 0.005         | < 0.03                  |
| Scalar     | 1674.537 (219) <sup>***</sup> | 0.843 | 0.040 | 0.034 | 0.068        | < 0.01                 | 0.008          | < 0.015                  | 0.007         | < 0.01                  |

Note: These analyses test for measurement invariance with cultures grouped into one of three categories: Western (such as US and Germany), “sometimes Western” (such as Russia), and non-Western (such as China and India). [Supplementary Section 24](#) defines this grouping in more detail. These tests account for the fact that individual participants are nested in groups.

**Configural:** Configural invariance tests whether the same factor structure fits across cultures (for example, taking care of elderly parents is a part of responsibilism in different cultures).

**Metric:** Metric invariance adds an extra layer of invariance by testing whether the factor loadings (the direction and strength of the relationships between individual questions and the underlying responsibilism factors) are the same across cultures.

**Scalar:** Scalar invariance adds an extra layer of invariance beyond metric invariance by testing whether cultures with the same underlying responsibilism scores also have the same average scores for individual questions on the responsibilism scale (item intercepts).

$\Delta$  = Change in fit indexes compared to the model above.

**Judging configural fit:** Readers can judge the fit of the configural models against recommended cutoffs for the fit indexes of > 0.90 for CFI, < 0.08 for RMSEA, and < 0.08 for SRMR.

**Judging metric fit:** Readers can judge the fit of the metric model by looking at the change ( $\Delta$ ) in model fit compared to the configural model. One suggestion for cutoffs is that  $\Delta$ CFI < 0.02,  $\Delta$ RMSEA < 0.015, and  $\Delta$ SRMR < 0.03<sup>72,73</sup>.

**CFI** = Comparative fit index.

**RMSEA** = Root mean squared error of approximation.

**SRMR** = Standardized root mean squared residual.

<sup>\*\*\*</sup>  $P < 0.001$



**Table S18B. Fit indices for measurement invariance across 100 cultures.**

| Model      | $\chi^2(DF)$      | CFI   | RMSEA | SRMR  | $\Delta CFI$ | $\Delta CFI$<br>Cutoff | $\Delta RMSEA$ | $\Delta RMSEA$<br>Cutoff | $\Delta SRMR$ | $\Delta SRMR$<br>Cutoff |
|------------|-------------------|-------|-------|-------|--------------|------------------------|----------------|--------------------------|---------------|-------------------------|
| Configural | 446.427 (106)***  | 0.938 | 0.016 | 0.016 |              |                        |                |                          |               |                         |
| Metric     | 643.598 (134)***  | 0.908 | 0.018 | 0.017 | 0.030        | 0.020                  | 0.002          | < 0.015                  | 0.001         | < 0.03                  |
| Scalar     | 3536.606 (131)*** | 0.382 | 0.046 | 0.023 | 0.526        | 0.068                  | 0.028          | < 0.015                  | 0.006         | < 0.01                  |

Note: These analyses test for measurement invariance across 100 cultural groups. Some researchers have argued that it is unreasonable to expect invariance<sup>86</sup> with such a large sample of cultures because it becomes more and more difficult to reach invariance as the number of cultures increases. These tests account for the fact that individual participants are nested in cultures.

5 **Configural:** Configural invariance tests whether the same factor structure fits across cultures (for example, taking care of elderly parents is a part of responsibilism in different cultures).

**Metric:** Metric invariance adds an extra layer of invariance by testing whether the factor loadings (the direction and strength of the relationships between individual questions and the underlying responsibilism factors) are the same across cultures.

10 **Scalar:** Scalar invariance adds an extra layer of invariance beyond metric invariance by testing whether cultures with the same underlying responsibilism scores also have the same average scores for individual questions on the responsibilism scale (item intercepts).

$\Delta$  = Change in fit indexes compared to the model above.

**Judging configural fit:** Readers can judge the fit of the configural models against recommended cutoffs for the fit indexes of > 0.90 for CFI, < 0.08 for RMSEA, and < 0.08 for SRMR.

15 **Judging metric fit:** Readers can judge the fit of the metric model by looking at the change ( $\Delta$ ) in model fit compared to the configural model. One suggestion for cutoffs is that  $\Delta CFI < 0.02$ ,  $\Delta RMSEA < 0.015$ , and  $\Delta SRMR < 0.03$ <sup>72,73</sup>. The CFI cutoff comes from recommendations specific to tests for large groups<sup>73</sup>.

**CFI** = Comparative fit index.

**RMSEA** = Root mean squared error of approximation.

**SRMR** = Standardized root mean squared residual.

20 \*\*\* $P < 0.001$

**Table S19A. Culture-level scale reliability.**

|                  | <b>Corrected Item-Total Correlation</b> | <b>Omega if Item Deleted</b> |
|------------------|-----------------------------------------|------------------------------|
| Item 1           | 0.525                                   | 0.82                         |
| Item 2           | 0.314                                   | 0.83                         |
| Item 2b Positive | 0.400                                   | 0.83                         |
| Item 3 Positive  | 0.435                                   | 0.82                         |
| Item 4 Positive  | 0.619                                   | 0.81                         |
| Item 5           | 0.716                                   | 0.80                         |
| Item 6           | 0.701                                   | 0.81                         |
| Item 7 Positive  | 0.330                                   | 0.83                         |
| Item 8           | 0.565                                   | 0.82                         |
| Item 9 Positive  | 0.271                                   | 0.83                         |
| Item 10 Positive | 0.483                                   | 0.82                         |
| Item 11 Positive | 0.423                                   | 0.83                         |
| Item 11b         | 0.040                                   | 0.84                         |
| Item 12 Positive | 0.432                                   | 0.83                         |
| Item 13 Positive | 0.329                                   | 0.83                         |
| Item 14 Positive | 0.383                                   | 0.83                         |

Weighted Reliability<sup>†</sup> = 0.89, Omega = 0.83

Note: Researchers have suggested cutoffs of 0.50 and 0.60 for reliability<sup>83,84</sup>. Values above the cutoff mean that the scale items are highly correlated with each other and seem to be measuring the same underlying concept. Our approach was to ask about different relationships and contexts, which has the upside of capturing a wider range of collectivism but will not produce the highest reliability values. Although we report reliability values, we believe external validity is more important than correlations between scale items. Internal validity (high correlations of the items) is not required for external validity. [Supplementary Section 1](#) reports tests of external validity.

“Positive”: Items marked “positive” are reverse coded so that higher values represent higher responsibility.

<sup>†</sup>Weighted reliability is also called “coefficient h.” This measure of reliability uses different weights for different scale items, similar to how structural equation models estimate responsibility scores using different weights for different items, as in [Fig. S16](#).

**Table S19B. Exploratory factor analysis loadings.**

|                       | Rotated Component Loadings |               |              |               |              |
|-----------------------|----------------------------|---------------|--------------|---------------|--------------|
|                       | 1                          | 2             | 3            | 4             | 5            |
| Item 1                | <b>0.861</b>               | -0.050        | 0.043        | -0.118        | 0.084        |
| Item 2D <sup>a</sup>  | -0.060                     | <b>0.858</b>  | -0.080       | 0.046         | 0.158        |
| Item 3                | 0.096                      | <b>-0.510</b> | 0.215        | -0.195        | 0.259        |
| Item 4                | -0.072                     | <b>-0.748</b> | 0.138        | 0.087         | -0.016       |
| Item 5                | <b>0.589</b>               | 0.370         | -0.164       | 0.045         | -0.136       |
| Item 6                | <b>0.872</b>               | -0.032        | 0.046        | -0.040        | -0.122       |
| Item 7                | 0.201                      | -0.251        | -0.243       | 0.223         | <b>0.697</b> |
| Item 8                | 0.461                      | <b>0.561</b>  | 0.351        | -0.138        | 0.094        |
| Item 9                | 0.116                      | -0.150        | <b>0.894</b> | 0.088         | -0.257       |
| Item 10               | -0.280                     | -0.132        | <b>0.652</b> | -0.273        | 0.027        |
| Item 11D <sup>a</sup> | 0.343                      | -0.252        | -0.167       | <b>-0.777</b> | 0.085        |
| Item 12               | 0.129                      | 0.020         | <b>0.729</b> | 0.441         | 0.166        |
| Item 13               | 0.072                      | -0.166        | 0.036        | <b>0.831</b>  | 0.095        |
| Item 14               | -0.229                     | 0.295         | -0.006       | -0.097        | <b>0.845</b> |
| Eigenvalues           | 3.832                      | 1.930         | 1.786        | 1.251         | 1.166        |

Note: This table reports the factor loadings from a principal component analysis using promax rotation with Kaiser Normalization. The largest loadings for each item **are in bold**. We prioritized the culture-level dataset to do the exploratory factor analysis because we were more interested in the factor structure at the culture level. The five factors accounted for 71.2% of the total variance. <sup>a</sup>The “D” marks paired items, where we take the difference between one item and the paired item. These items can help detect acquiescent response styles. To determine the final factor structure ([Fig. S16](#)), we took into consideration the exploratory factor analysis loadings, measurement invariance results, and theoretical and practical considerations. We describe these in more detail in Supplementary Sections 6 and 7.

**Table S20. Individual-level scale reliability.**

|                  | Corrected Item-Total Correlation | Omega if Item Deleted |
|------------------|----------------------------------|-----------------------|
| Item 1           | .284                             | 0.49                  |
| Item 2           | .252                             | 0.49                  |
| Item 2b Positive | .165                             | 0.51                  |
| Item 3 Positive  | .134                             | 0.51                  |
| Item 4 Positive  | .152                             | 0.51                  |
| Item 5           | .345                             | 0.46                  |
| Item 6           | .341                             | 0.47                  |
| Item 7 Positive* | -.017                            | 0.54                  |
| Item 8           | .225                             | 0.50                  |
| Item 9 Positive  | .133                             | 0.51                  |
| Item 10 Positive | .216                             | 0.49                  |
| Item 11 Positive | .233                             | 0.50                  |
| Item 11b         | .071                             | 0.53                  |
| Item 12 Positive | .125                             | 0.52                  |
| Item 13 Positive | .081                             | 0.53                  |
| Item 14 Positive | .060                             | 0.53                  |

Weighted Reliability<sup>†</sup> = 0.63, Omega = 0.52

Note: Researchers have suggested cutoffs of 0.50 and 0.60 for reliability<sup>83,84</sup>. Values above the cutoff mean that the scale items are highly correlated with each other and seem to be measuring the same underlying concept. Our approach was to ask about different relationships and contexts, which has the upside of capturing a broader slice of collectivism but will not produce the highest reliability values. Although we report reliability values, we believe external validity is more important than correlations between scale items. Internal validity (high correlations of the items) is not required for external validity. [Supplementary Section 1](#) reports tests of external validity.

“Positive”: Items marked “positive” are reverse coded so that higher values represent higher responsibility.

\*Item 7 had a low item-total correlation at the individual level (although higher at the culture level). Thus, we tested whether removing it would improve external validity or sensitivity in measuring cultural differences. Removing Item 7 decreased correlations with external markers of collectivism calculated at the culture level ( $r = 0.675$  with all items and  $r = 0.661$  after removing Item 7). Removing Item 7 also decreased the sensitivity of the scale in detecting differences individual differences within cultures. For example, in the US sample, removing Item 7 decreased the correlation with political conservatism from  $r = 0.230$  to  $r = 0.201$ . Based on these results, we chose to keep Item 7.

<sup>†</sup>Weighted reliability is also called “coefficient h.” This measure of reliability uses different weights for different scale items, similar to how structural equation models estimate responsibility scores using different weights for different items, as in [Fig. S16](#).

**Table S21. Recruitment methods by culture.**

| <b>Culture</b> | <b>Recruitment</b> | <b>Culture</b>  | <b>Recruitment</b>                         | <b>Culture</b>  | <b>Recruitment</b>               |
|----------------|--------------------|-----------------|--------------------------------------------|-----------------|----------------------------------|
| Algeria        | Survey Comp.       | India           | Collaborator                               | Peru            | Collaborator                     |
| Angola         | Collaborator       | Indonesia       | Collaborator                               | Philippines     | Collaborator                     |
| Argentina      | Survey Comp.       | Iran            | Collaborator                               | Poland          | Prolific                         |
| Australia      | Collaborator       | Iraq            | Survey Comp.                               | Portugal        | Collaborator                     |
| Austria        | Prolific           | Ireland         | Collaborator                               | Puerto Rico     | Collaborator                     |
| Bahamas        | Collaborator       | Israel          | Collaborator                               | Romania         | Collaborator                     |
| Bangladesh     | Collaborator       | Italy           | Collaborator                               | Russia          | Collaborator                     |
| Barbados       | Collaborator       | Ivory Coast     | Collaborator                               | Saudi Arabia    | Collaborator                     |
| Belarus        | Collaborator       | Japan           | Collaborator                               | Senegal         | Collaborator                     |
| Belgium        | Collaborator       | Kazakhstan      | Collaborator                               | Serbia          | Collaborator                     |
| Brazil         | Collaborator       | Kenya           | Survey Comp.,<br>Collaborator              | Singapore       | Collaborator                     |
| Canada         | Collaborator       | Kuwait          | Collaborator                               | Slovakia        | Collaborator                     |
| Chile          | Collaborator       | Latvia          | Prolific                                   | Slovenia        | Prolific                         |
| China          | Collaborator       | Lebanon         | Collaborator                               | South Africa    | Collaborator                     |
| Colombia       | Collaborator       | Macau           | Collaborator                               | South Korea     | Collaborator                     |
| Costa Rica     | Collaborator       | Madagascar      | Collaborator                               | Spain           | Collaborator                     |
| Croatia        | Collaborator       | Malawi          | Collaborator                               | Sri Lanka       | Collaborator                     |
| Cyprus         | Collaborator       | Malaysia        | Collaborator                               | Sweden          | Prolific                         |
| Czech Rep.     | Prolific           | Mali            | Collaborator                               | Switzerland     | Prolific                         |
| Denmark        | Collaborator       | Mexico          | Collaborator                               | Taiwan          | Collaborator                     |
| DRC            | Collaborator       | Mongolia        | Collaborator                               | Thailand        | Collaborator                     |
| Egypt          | Collaborator       | Morocco         | Collaborator                               | Trinidad & Tob. | Collaborator                     |
| Estonia        | Prolific           | Mozambique      | Collaborator                               | Turkey          | Collaborator                     |
| Ethiopia       | Survey Comp.       | Myanmar         | Survey Comp.,<br>Collaborator              | Uganda          | Collaborator                     |
| Finland        | Prolific           | Namibia         | Collaborator                               | UK              | Collaborator,<br>Prolific        |
| France         | Collaborator       | Nepal           | Collaborator                               | Ukraine         | Collaborator                     |
| Germany        | Collaborator       | Netherlands     | Collaborator,<br>Survey Comp.,<br>Prolific | United States   | Collaborator,<br>Prolific, MTurk |
| Ghana          | Collaborator       | New Zealand     | Collaborator                               | Uzbekistan      | Collaborator                     |
| Greece         | Collaborator       | Nicaragua       | Survey Comp.                               | Venezuela       | Collaborator                     |
| Greenland      | Collaborator       | Nigeria         | Collaborator                               | Vietnam         | Collaborator                     |
| Guyana         | Collaborator       | North Macedonia | Collaborator                               | Zambia          | Collaborator                     |
| Hong Kong      | Collaborator       | Norway          | Collaborator                               | Zimbabwe        | Collaborator                     |
| Hungary        | Collaborator       | Oman            | Survey Comp.                               |                 |                                  |
| Iceland        | Collaborator       | Pakistan        | Collaborator                               |                 |                                  |

Note: Collaborators were people in the local culture (such as professors) who recruited mostly college students locally. We recruited survey company samples through the company Cint, which recruited college student samples. Prolific is a crowdwork platform similar to MTurk. On Prolific, we set restrictions to people born in the country and current college students.

**Table S22. Responsibilism scores by culture.**

| <b>Culture</b> | <b>Responsibilism</b> | <b>Culture</b> | <b>Respons.</b> | <b>Culture</b>  | <b>Respons.</b> |
|----------------|-----------------------|----------------|-----------------|-----------------|-----------------|
| Algeria        | 74.95                 | India          | 84.66           | Pakistan        | 82.76           |
| Angola         | 82.64                 | Indonesia      | 76.54           | Peru            | 52.70           |
| Argentina      | 59.59                 | Iran           | 80.75           | Philippines     | 77.22           |
| Australia      | 30.01                 | Iraq           | 88.49           | Poland          | 55.61           |
| Austria        | 34.57                 | Ireland        | 31.25           | Portugal        | 47.21           |
| Bahamas        | 69.80                 | Israel         | 56.38           | Puerto Rico     | 45.74           |
| Bangladesh     | 70.17                 | Italy          | 47.73           | Romania         | 35.68           |
| Barbados       | 72.49                 | Ivory Coast    | 83.33           | Russia          | 58.04           |
| Belarus        | 60.62                 | Japan          | 56.51           | Saudi Arabia    | 69.38           |
| Belgium        | 28.45                 | Kazakhstan     | 56.38           | Senegal         | 88.81           |
| Brazil         | 56.01                 | Kenya          | 84.75           | Serbia          | 49.82           |
| Canada         | 32.28                 | Kuwait         | 100.00          | Singapore       | 69.04           |
| Chile          | 52.63                 | Latvia         | 59.07           | Slovakia        | 55.85           |
| China          | 93.92                 | Lebanon        | 72.74           | Slovenia        | 40.70           |
| Colombia       | 62.06                 | Macau          | 81.60           | South Africa    | 73.22           |
| Costa Rica     | 62.98                 | Macedonia      | 46.82           | South Korea     | 64.12           |
| Croatia        | 46.99                 | Madagascar     | 89.74           | Spain           | 59.49           |
| Cyprus         | 45.62                 | Malawi         | 71.38           | Sri Lanka       | 81.71           |
| Czech Rep.     | 59.58                 | Malaysia       | 72.00           | Sweden          | 31.05           |
| Denmark        | 0.00*                 | Mali           | 82.06           | Switzerland     | 30.03           |
| DRC            | 71.71                 | Mexico         | 62.06           | Taiwan          | 90.48           |
| Egypt          | 88.44                 | Mongolia       | 63.29           | Thailand        | 75.38           |
| Estonia        | 33.74                 | Morocco        | 84.09           | Trinidad & Tob. | 68.79           |
| Ethiopia       | 77.60                 | Mozambique     | 82.95           | Turkey          | 77.25           |
| Finland        | 53.60                 | Myanmar        | 72.97           | Uganda          | 80.60           |
| France         | 25.48                 | Namibia        | 87.01           | UK              | 36.66           |
| Germany        | 16.79                 | Nepal          | 71.91           | Ukraine         | 56.88           |
| Ghana          | 79.65                 | Netherlands    | 20.41           | United States   | 40.88           |
| Greece         | 35.80                 | New Zeal.      | 22.37           | Uzbekistan      | 61.77           |
| Greenland      | 55.77                 | Nicaragua      | 71.38           | Venezuela       | 62.01           |
| Guyana         | 78.38                 | Nigeria        | 75.67           | Vietnam         | 85.72           |
| Hong Kong      | 83.58                 | Norway         | 14.95           | Zambia          | 87.39           |
| Hungary        | 36.22                 | Oman           | 81.38           | Zimbabwe        | 74.51           |
| Iceland        | 25.51                 |                |                 |                 |                 |

Note: We scaled the scores so that the lowest culture score is zero, and the highest culture score is 100.

However, it is possible for future samples to score above 100 or below zero. \*Denmark is the culture with the lowest score, so we used this as the zero point for culture scores.

**Table S23. Sample demographics excluding non-locals and careless responders (part 1).**

| <b>Culture</b> | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b> | <b>Site</b>        | <b>Compensation</b>      |
|----------------|----------|---------------|------------|-----------|-----------------|-----------------|--------------------|--------------------------|
| Algeria        | 55       | Stud./Adult   | 30.50      | 9.70      | 37.0            | Arabic          | Various            | Payment                  |
| Angola         | 164      | Student       | 25.22      | 5.50      | 51.8            | Portuguese      | Luanda             | Payment                  |
| Argentina      | 85       | Student       | 19.83      | 1.48      | 56.0            | Spanish         | Various            | Payment                  |
| Australia      | 74       | Student       | 24.93      | 8.77      | 84.3            | English         | Perth              | Course Credit            |
| Austria        | 66       | Stud./Adult   | 27.79      | 6.56      | 45.5            | German          | Various            | Payment                  |
| Bahamas        | 22       | Student       | 27.00      | 8.98      | 63.6            | English         | Various            | Prize Raffle             |
| Bangladesh     | 98       | Student       | 22.64      | 1.65      | 48.0            | Bengali         | Dhaka              | Payment, CC <sup>a</sup> |
| Barbados       | 72       | Student       | 26.89      | 9.59      | 69.0            | English         | Bridgetown         | Prize Raffle             |
| Belarus        | 59       | Stud./Adult   | 27.22      | 9.85      | 55.9            | Russian         | Minsk              | Payment                  |
| Belgium        | 63       | Student       | 21.60      | 5.89      | 76.2            | French          | Louvain-la-Neuve   | Prize Raffle             |
| Brazil         | 98       | Student       | 23.48      | 6.75      | 67.3            | Portuguese      | Aracaju            | Course Credit            |
| Canada         | 103      | Student       | 21.55      | 5.39      | 81.6            | English         | Waterloo           | Course Credit            |
| Chile          | 98       | Student       | 22.28      | 8.32      | 72.0            | Spanish         | Santiago           | Payment                  |
| China          | 207      | Student       | 18.42      | 1.11      | 71.1            | Simpl. Chinese  | Wuhan              | Course Credit            |
| Colombia       | 65       | Student       | 22.55      | 6.13      | 78.5            | Spanish         | Cali               | Payment                  |
| Costa Rica     | 29       | Student       | 23.52      | 3.02      | 65.5            | Spanish         | San Jose           | Payment                  |
| Croatia        | 88       | Student       | 22.40      | 3.36      | 84.1            | Croatian        | Zagreb             | Payment                  |
| Cyprus         | 114      | Student       | 24.66      | 5.33      | 62.3            | Greek           | Nicosia            | Payment                  |
| Czech Rep.     | 98       | Stud./Adult   | 25.18      | 4.99      | 37.8            | Czech           | Various            | Payment                  |
| Denmark        | 128      | Student       | 24.45      | 5.28      | 80.5            | English         | Copenhagen, Aarhus | Course Credit            |
| DRC            | 45       | Student       | 24.32      | 3.54      | 8.9             | French          | Lubumbashi         | Payment                  |
| Egypt          | 683      | Student       | 20.00      | 2.17      | 84.2            | Arabic          | Menofia            | Payment                  |
| Estonia        | 74       | Student       | 23.79      | 4.75      | 55.4            | Estonian        | Various            | Payment                  |
| Ethiopia       | 41       | Student       | 21.44      | 3.65      | 56.1            | Amharic, Oromo  | Various            | Payment                  |
| Finland        | 92       | Stud./Adult   | 28.78      | 8.62      | 32.6            | Finnish         | Various            | Payment                  |
| France         | 82       | Stud./Adult   | 28.20      | 11.07     | 72.0            | French          | Occitanie, Paris   | Volun., Raffle           |
| Germany        | 49       | Student       | 21.06      | 2.46      | 83.7            | German          | Mannheim           | Course Credit            |
| Ghana          | 129      | Student       | 20.88      | 3.72      | 70.3            | English         | Accra              | Payment                  |
| Greece         | 146      | Student       | 26.29      | 10.43     | 66.2            | Greek           | Athens             | Payment                  |
| Greenland      | 54       | Student       | 29.83      | 7.05      | 86.5            | Greenlandic     | Nuuk               | Prize Raffle             |
| Guyana         | 186      | Student       | 29.69      | 7.91      | 88.7            | English         | Georgetown         | Prize Raffle             |
| Hong Kong      | 305      | Student       | 22.97      | 5.82      | 79.0            | English         | Hong Kong          | Payment                  |
| Hungary        | 110      | Student       | 26.60      | 9.00      | 82.7            | Hungarian       | Pécs               | Volunteer                |
| Iceland        | 108      | Student       | 25.19      | 6.95      | 58.3            | Icelandic       | Reykjavik          | Prize Raffle             |
| India          | 84       | Student       | 24.10      | 6.80      | 42.2            | English         | Chandigarh         | Payment                  |

<sup>a</sup>CC = Course credit. “Various” = internet samples, where participants lived in many places.



**Table S24. Sample demographics excluding non-locals and careless responders (part 2).**

| <b>Culture</b>  | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b> | <b>Site</b>            | <b>Compensation</b>      |
|-----------------|----------|---------------|------------|-----------|-----------------|-----------------|------------------------|--------------------------|
| Indonesia       | 319      | Student       | 20.30      | 2.50      | 72.0            | Indonesian      | Pekanbaru              | Payment                  |
| Iran            | 73       | Student       | 29.59      | 5.27      | 57.5            | Persian         | Tehran                 | Volunteer                |
| Iraq            | 36       | Student       | 21.11      | 2.94      | 33.3            | Arabic          | Various                | Payment                  |
| Ireland         | 79       | Student       | 19.29      | 1.43      | 78.5            | English         | Limerick               | Volunteer                |
| Israel          | 99       | Student       | 25.44      | 2.69      | 48.5            | Hebrew          | Various                | Payment                  |
| Italy           | 64       | Student       | 28.84      | 9.82      | 81.3            | Italian         | Naples                 | Volunteer                |
| Ivory Coast     | 20       | Student       | 21.95      | 6.48      | 35.0            | French          | Abidjan                | Payment                  |
| Japan           | 160      | Student       | 20.03      | 3.37      | 46.8            | Japanese        | Hokkaido, Tokyo        | Course Credit            |
| Kazakhstan      | 191      | Stud./Adult   | 27.04      | 8.91      | 52.6            | Russian         | Pavlodar               | Course Credit            |
| Kenya           | 315      | Student       | 22.55      | 2.85      | 46.3            | English         | Nairobi, Various       | Payment                  |
| Kuwait          | 54       | Student       | 24.06      | 5.65      | 74.1            | Arabic          | Kuwait                 | Course Credit            |
| Latvia          | 54       | Student       | 20.91      | 2.15      | 42.6            | Latvian         | Various                | Payment                  |
| Lebanon         | 78       | Student       | 18.65      | 1.07      | 57.7            | English         | Beirut                 | Volunteer                |
| Macau           | 65       | Student       | 24.25      | 5.69      | 67.7            | Trad. Chinese   | Macau                  | Payment                  |
| Madagascar      | 98       | Student       | 20.42      | 2.39      | 50.0            | French          | Antananarivo           | Payment                  |
| Malawi          | 61       | Student       | 22.62      | 2.65      | 36.1            | English         | Blantyre               | Payment                  |
| Malaysia        | 62       | Student       | 25.74      | 6.67      | 66.1            | English, Malay  | Kuala Lumpur           | Payment                  |
| Mali            | 56       | Student       | 21.66      | 2.06      | 30.4            | French          | Bamako                 | Payment                  |
| Mexico          | 107      | Student       | 19.81      | 1.40      | 60.4            | Spanish         | Mex. City, Guadalajara | Course Credit            |
| Mongolia        | 125      | Student       | 23.26      | 7.04      | 64.8            | Mongolian       | Ulaanbaatar            | Volunteer                |
| Morocco         | 44       | Student       | 18.59      | 0.73      | 84.1            | French          | Tangier                | Course Credit            |
| Mozambique      | 53       | Student       | 20.62      | 2.27      | 34.0            | Portuguese      | Maputo                 | Payment                  |
| Myanmar         | 92       | Student       | 21.54      | 3.39      | 59.8            | Burmese         | Magway, Various        | Payment                  |
| Namibia         | 112      | Student       | 23.96      | 5.51      | 57.1            | English         | Kavango                | Payment                  |
| Nepal           | 118      | Student       | 29.58      | 9.61      | 63.2            | Nepali          | Kathmandu              | Payment                  |
| Netherlands     | 225      | Student       | 21.16      | 4.01      | 63.1            | Dutch           | Amsterdam, Various     | Payment, CC <sup>a</sup> |
| New Zealand     | 72       | Student       | 22.16      | 4.71      | 45.1            | English         | Wellington             | Payment                  |
| Nicaragua       | 48       | Student       | 20.77      | 2.04      | 50.0            | Spanish         | Managua                | Payment                  |
| Nigeria         | 48       | Student       | 20.13      | 1.92      | 68.8            | English         | Lagos                  | Payment                  |
| North Macedonia | 50       | Stud./Adult   | 32.52      | 8.82      | 74.0            | Macedonian      | Tetovo                 | Prize Raffle             |
| Norway          | 118      | Student       | 23.53      | 5.28      | 76.3            | Norweg., Eng.   | Oslo                   | Course Credit            |
| Oman            | 36       | Student       | 20.17      | 1.81      | 44.4            | Arabic          | Various                | Payment                  |
| Pakistan        | 341      | Student       | 23.66      | 7.09      | 50.8            | English         | Haripur                | Donation to Charity      |
| Peru            | 79       | Student       | 27.16      | 9.63      | 60.0            | Spanish         | Lima                   | Volunteer                |
| Philippines     | 61       | Student       | 24.18      | 6.57      | 84.7            | English         | Cebu                   | Volunteer                |

<sup>a</sup>CC = Course credit. “Various” = internet samples, where participants lived in many places.

**Table S25. Sample demographics excluding non-locals and careless responders (part 3).**

| <b>Culture</b>       | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b>               | <b>Site</b>                                                                       | <b>Compensation</b>         |
|----------------------|----------|---------------|------------|-----------|-----------------|-------------------------------|-----------------------------------------------------------------------------------|-----------------------------|
| Poland               | 89       | Student       | 21.70      | 3.66      | 28.1            | Polish                        | Various                                                                           | Payment                     |
| Portugal             | 86       | Student       | 26.01      | 8.29      | 75.6            | Portuguese                    | Lisbon                                                                            | Prize Raffle                |
| Puerto Rico          | 66       | Student       | 21.74      | 2.49      | 74.2            | Spanish                       | Various                                                                           | Payment                     |
| Romania              | 118      | Student       | 24.21      | 7.61      | 83.8            | Romanian                      | Botosani                                                                          | Payment                     |
| Russia               | 166      | Stud./Adult   | 29.15      | 12.92     | 57.0            | Russian                       | Moscow                                                                            | CC <sup>a</sup> , Volunteer |
| Saudi Arabia         | 104      | Student       | 24.11      | 7.22      | 63.4            | Arabic                        | Tabuk                                                                             | Volunteer                   |
| Senegal              | 382      | Student       | 21.95      | 2.81      | 61.2            | French                        | Dakar                                                                             | Course Credit               |
| Serbia               | 89       | Stud./Adult   | 35.31      | 11.26     | 84.1            | Serbian                       | Various                                                                           | Volunteer                   |
| Singapore            | 50       | Student       | 21.10      | 1.57      | 68.0            | English                       | Singapore                                                                         | Course Credit               |
| Slovakia             | 115      | Student       | 22.36      | 2.65      | 67.8            | Slovak                        | Various                                                                           | Payment                     |
| Slovenia             | 96       | Stud./Adult   | 25.79      | 8.70      | 27.1            | Slovenian                     | Various                                                                           | Payment                     |
| South Africa         | 70       | Student       | 22.88      | 3.14      | 64.7            | English                       | Johannesburg,<br>Pretoria                                                         | Volunteer                   |
| South Korea          | 73       | Student       | 21.04      | 2.61      | 58.9            | Korean                        | Seoul                                                                             | Course Credit               |
| Spain                | 48       | Student       | 22.67      | 1.51      | 50.0            | Spanish                       | Madrid                                                                            | Volunteer                   |
| Sri Lanka            | 92       | Student       | 23.30      | 1.38      | 80.4            | Sinhala                       | Colombo                                                                           | Payment                     |
| Sweden               | 44       | Student       | 25.75      | 6.52      | 40.9            | Swedish                       | Various                                                                           | Payment                     |
| Switzerland          | 32       | Student       | 26.38      | 6.20      | 37.5            | German,<br>French,<br>Italian | Various                                                                           | Payment                     |
| Taiwan               | 165      | Stud./Adult   | 37.61      | 8.94      | 61.2            | Trad.<br>Chinese              | Taipei                                                                            | Volunteer                   |
| Thailand             | 301      | Student       | 20.51      | 2.41      | 77.4            | Thai                          | Udon Thani,<br>Bangkok                                                            | Payment                     |
| Trinidad &<br>Tobago | 51       | Student       | 21.16      | 2.03      | 82.4            | English                       | San Juan                                                                          | Payment                     |
| Turkey               | 141      | Student       | 21.78      | 1.39      | 52.5            | Turkish                       | Istanbul                                                                          | CC <sup>a</sup> , Volunteer |
| Uganda               | 32       | Student       | 21.59      | 1.50      | 71.9            | English                       | Kampala                                                                           | Payment                     |
| UK                   | 441      | Student       | 29.90      | 10.55     | 59.4            | English                       | London, Various                                                                   | Payment                     |
| Ukraine              | 181      | Stud./Adult   | 33.22      | 13.69     | 75.7            | Ukrainian                     | Kyiv                                                                              | CC <sup>a</sup> , Volunteer |
| United States        | 511      | Student       | 20.07      | 3.60      | 63.4            | English                       | Ann Arbor, Greenville,<br>LA, Richardson,<br>Charlottesville, Chicago,<br>Various | Payment,<br>Course Credit   |
| Uzbekistan           | 82       | Student       | 21.83      | 1.69      | 69.1            | Russian                       | Tashkent                                                                          | Volunteer                   |
| Venezuela            | 56       | Student       | 20.43      | 1.40      | 58.9            | Spanish                       | Various                                                                           | Payment                     |
| Vietnam              | 126      | Student       | 21.94      | 3.51      | 83.3            | Vietnamese                    | Ho Chi Minh                                                                       | Payment                     |
| Zambia               | 74       | Student       | 25.19      | 6.72      | 51.4            | English                       | Lusaka                                                                            | Payment                     |
| Zimbabwe             | 779      | Student       | 31.89      | 7.51      | 50.5            | English                       | Harare                                                                            | Payment                     |

<sup>a</sup>CC = Course credit. “Various” = internet samples, where participants lived in many places.

**Table S26. Initial responsibilism scale before item selection based on validity and reliability testing.**

|     | Item                                                                                                                                                                                                                                                                                                                                                                                              | Rationale for Creating the Item                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1   | Children should be taught to place duty before pleasure.                                                                                                                                                                                                                                                                                                                                          | Responsibilistic cultures emphasize duties in close relationships. The use of “children” implies that this is within the family. This item comes from Singelis and colleagues’ 1995 scale <sup>3</sup> , which demonstrates that the earlier scales contained some items that tapped into the responsibilistic vision of collectivism.                                                                                                                                                                                  |
| 2A  | You’re walking on the street, and a stranger starts talking to you.<br><br>The stranger is probably trying to trick me.                                                                                                                                                                                                                                                                           | Responsibilistic cultures have stronger duties to close others but weaker duties and trust toward strangers. This item avoids vague wordings by specifying that the other person is a stranger and that the context is walking on the street.                                                                                                                                                                                                                                                                           |
| 2B* | I would trust the stranger. The person probably has good intentions.                                                                                                                                                                                                                                                                                                                              | Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items. This can reveal cultural differences in how people answer surveys.                                                                                                                                                                                                                                                                                                                               |
| 3*  | It’s after dinner, and you go to the ATM to get some cash. The person in front of you turns around and says, “I can’t figure out how to get the machine to take a deposit. Can you help me do it?”<br><br>I would trust the person.                                                                                                                                                               | Responsibilistic cultures have stronger duties to close others but weaker duties and trust toward strangers. This item avoids abstract, unclear wordings by specifying that the other person is a stranger and that the context is evening and at an ATM. We anticipate needing to update this item as cashless payments become more common.                                                                                                                                                                            |
| 4†  | Imagine you are married and have a child. Your child wants to take extracurricular classes on the piano, but the classes are very expensive. You can afford the classes if you use all of your meager retirement savings, but you are nearing retirement age, and you may need that money soon.<br><br>As a parent, I have the responsibility to sacrifice myself for my child in this situation. | This item asks about parents sacrificing for their children. The item details specific close relationships and explains exactly what the sacrifice is to avoid vagueness. However, this item did not pass checks of equivalence across cultures and was removed from the final version. This item may have failed because extracurricular activities are very important for advancement in some cultures (such as the US) but not in others (such as in China, where the college entrance exam is extremely important). |

5\*

You are working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You've been working on a sales account that could be a big accomplishment. The co-worker asks about the account.

I would feel comfortable telling this co-worker about the account.

People in responsibilistic cultures are more wary of non-close ties. This item specifically tests the idea that belonging to the same big group is *not* sufficient for tight ties in responsibilistic cultures. In this scenario, the other person is in the same group (the same company) but is unfamiliar. We theorized that people in responsibilistic cultures would trust this co-worker less because the tie is weak. This fits with prior research finding that people in collectivistic cultures are more vigilant toward people in large ingroups<sup>27</sup>.

6

You're working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You are sitting outside a meeting room waiting for the meeting to start, and the two of you start talking. The co-worker starts asking you if you're married, if you have children, what your hobbies are. Then you remember what a friend of yours told you: "Be careful not to share your personal secrets with people you don't know because they could use those secrets to hurt you."

**Do you agree with this statement?**

People in responsibilistic cultures are more wary of non-close ties. This item specifically tests the idea that belonging to the same big group is *not* sufficient for tight ties in responsibilistic cultures. In this scenario, the other person is in the same group (the same company) but is unfamiliar. We theorized that people in responsibilistic cultures would trust this co-worker less because the tie is weak. There is evidence that people in collectivistic cultures are more wary toward ingroup members from research on "enemyship" and vigilance comparing the US to Ghana, Mainland China, and Hong Kong<sup>27,64,103</sup>.

7

We should keep our aging parents with us at home.

Responsibilistic cultures emphasize duties in close relationships. The use of "parents" and "home" imply helping within the family. This sort of help is pragmatic and objective. It does not ask about warm feelings. This item comes from Singelis and colleagues' 1995 scale<sup>3</sup>, which demonstrates that the earlier scales contained some items that tapped into the responsibilistic vision of collectivism.

8\*

You start working at a new company. The salary is good, and you like the job. But when you get your first paycheck, you find out that you got more money than you expected. You look at the paycheck, and you see that they didn't take out the legally required taxes. You ask your boss, and your boss says, "Oh, everyone here does that. That's no problem. And this way you earn more money!"

I would insist on paying the required taxes.

Responsibilistic cultures have stronger duties in close relationships but weaker ties in distant relationships and impersonal systems. In the case of cheating on taxes, the harm is mostly to strangers—an impersonal system. The fact that everyone in the company evades this tax also suggests (a) that evading taxes would conform to the social norm and (b) breaking the rules would not likely bring punishment.

9\*†

You walk into an elevator, and a person you don't know enters at the same time.

As I wait to get to my floor, I would feel natural starting up a conversation with that person.

Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority. It may even be strange to speak to strangers without some obvious connection to that person. This item did not pass checks of equivalence across cultures and was removed from the final version. One potential explanation is that collaborators in some cultures explained that elevators are uncommon in their environment.

10A†

A good friend calls you on the phone, and says, "I'm in a hurry, and I need you to loan me some money." But your friend doesn't tell you the reason.

I would give them money right away.

Responsibilistic cultures have stronger duties in close relationships. However, this item did not pass checks of equivalence across cultures and was removed from the final version. Collaborators in Malaysia and Indonesia explained that it is common for scammers to call and pretend to be a relative and ask for money, whereas these types of calls are rare in some cultures. This probably created differences in how people responded to this item across cultures that are unrelated to responsibilism.

10B\*†

I would ask my friend to explain why he/she needs the money.

Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items. This can reveal cultural differences in how people answer surveys. However, this item did not pass checks of equivalence across cultures and was removed from the final version.

|                   |                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 11A <sup>†</sup>  | <p>A good friend calls you and says, “I know a person who’s starting a company. If we each invest now, we’ll make 20% profit in two years.” So you both invest. After two years, you profit 20%, but you see your friend’s bank statement, and your friend profited 30% from the deal.</p> <p>My friend probably has a good reason for why there is the difference.</p> | <p>Responsibilistic cultures have stronger duties in close relationships and thus may give their close friends the benefit of the doubt. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One potential explanation is that the scenario many people do not have experience investing in companies, which makes the scenario hard to relate to.</p>                                                                                                                                                                                                                                                                                                                     |
| 11B* <sup>†</sup> | <p>I would ask my friend to explain why the numbers are different.</p>                                                                                                                                                                                                                                                                                                  | <p>Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items. This can reveal cultural differences in how people answer surveys. However, this item did not pass checks of equivalence across cultures and was removed from the final version.</p>                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 12                | <p>Children should feel honored if their parents receive a distinguished award.</p>                                                                                                                                                                                                                                                                                     | <p>Responsibilistic cultures emphasize duties in close relationships, which means that people’s accomplishments are less individually determined. Thus, children may feel more closely tied to their parents’ accomplishments and failures. The item specifies close relationships with the words “children” and “parents.” In addition, the description of the “distinguished award” implies social approval, similar to the concepts of honor and face, which tend to be stronger in collectivistic cultures<sup>11</sup>. This item comes from Singelis and colleagues’ 1995 scale<sup>3</sup>, which demonstrates that the earlier scales contained some items that tapped into the responsibilistic vision of collectivism.</p> |
| 13* <sup>†</sup>  | <p>If my parents have a nice little thing that happened to them today (for example, they saw a very pretty flower or their co-worker complimented them), my parents would tell me the little story.</p>                                                                                                                                                                 | <p>This item describes warm fuzzies—sharing little stories that have little pragmatic use but bring emotional closeness. Thus, we hypothesized that people in responsibilistic cultures would agree less with this item. However, this item did not pass checks of equivalence across cultures and was removed from the final version.</p>                                                                                                                                                                                                                                                                                                                                                                                           |

14\*

I like going to parties and talking to people I don't know in order to make new friends.

Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority. It may even be strange to speak to strangers without some obvious connection to that person. If relationships come with strong duties, it may even be overly costly to build too many relationships.

15\*

Imagine you're on a four-hour flight to another city, and you're traveling alone. Another person sits in the next seat to you.

I would talk to that person and share details about myself—my job, my family members, where I went to school, whether I'm married or have a boyfriend/girlfriend.

Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority. It may even be strange to speak to strangers without some obvious connection to that person. This item also taps into concerns about the potential for harm in building relationships too freely. The context of the plan and the details of the personal information keep the scenario concrete. The details prevent people from filling in the details differently across cultures. In some cultures, people said that taking domestic flights was not common and suggested replacing this with a train or bus ride.

16\*†

You're at a bar with your friends. When you're walking to the bathroom, you see a table of people you don't know, and you see that they're drinking your favorite kind of beer.

I would say something to them as I walk by. ("Hey, great choice!" Or "Hey, that's my favorite beer!")

Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority. It may even be odd to speak to strangers without some obvious connection to that person. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One possible explanation is that going to bars is uncommon in some cultures, such as Muslim-majority cultures that have strong prohibitions on alcohol.

17A†

You're at a bar with 10 very good friends. You came to celebrate your friend's birthday. Five minutes after you sit down, five people come over and say, "I'm sorry, we reserved this table. Could you move?" However, you clearly remember reserving this table. You think there was probably a mistake in the booking process.

I would suggest, "Oh, I'm sorry. We booked the table in advance. There are probably seats somewhere else."

Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority. Sharing a table in this situation may be odd in responsibilistic cultures because the other people are strangers. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One possible explanation is that going to bars is uncommon in some cultures. In addition, it may be uncommon to make reservations in some cultures.

|                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>17B*† I would suggest, “Oh, no problem, sit down with us. We can share!”</p>                                                                                                                                                                                                              | <p>Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items. This can reveal cultural differences in how people answer surveys. However, this item did not pass checks of equivalence across cultures and was removed from the final version.</p>                                                                                                                                                                                                                                                                                                   |
| <p>18*† Imagine you live in an apartment with four roommates that you didn’t know before you moved in. One day, two friends of your roommate come over. Your roommate says, “My room isn’t big enough for both of my friends. Can one sleep in your room?”</p> <p>I would happily agree.</p> | <p>This item asks about a friend of a roommate to test for a weak relationship. The word “happily” also implies warm feelings. We theorized that people in responsibilistic cultures would agree less with this item because it is about warm feelings with weak ties. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One possible explanation is that the tie to the roommate is still close enough to count as a close relationship. Another possibility is that this may be hard to relate to for students who live in shared dormitories.</p> |
| <p>19A You get a 60% on a midterm, and you might fail the class if you don’t do better on the final exam. You’re in a bad mood. Your good friend criticizes you. “It’s your fault. You need to study harder.” This makes you feel worse.</p> <p>This friend is being a good friend.</p>      | <p>This item contrasts warm fuzzies versus more long-term duties about people’s objective outcomes. In cultures where it is important to express warm feelings in a relationship, people should disagree with this item. But in cultures where people do not prioritize expressing positive feelings (and if they expect relationships to be binding enough that they can withstand harsh feedback), people should agree with this item.</p>                                                                                                                                                                     |
| <p>19B* I would try to make friends who are more supportive.</p>                                                                                                                                                                                                                             | <p>Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items. This can reveal cultural differences in how people answer surveys. Agreeing with this item also implies that people believe it is possible and acceptable to leave old friends and make new friends, which means that relationships are not strongly binding.</p>                                                                                                                                                                                                                      |



20\*<sup>†</sup> You're driving a car and you're waiting at a stoplight. It's summer, so the car windows are open. You hear a song you like playing in the car next to you.

I would shout to the other person, "Hey, nice song!"

Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority. It may even be odd to speak to strangers without some obvious connection to that person. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One explanation is that driving your own car is rare in many countries, especially for young people.

21A\*<sup>†</sup> You go to a yoga class for the first time at a yoga studio downtown. You don't know anyone in the class. After the class is over, you see a group of five people chatting. They all know each other.

I would stay and approach the group.

Responsibilistic cultures prioritize close relationships. Building ties with strangers is not usually a priority. It may even be odd to speak to strangers without some obvious connection to that person. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One explanation is that yoga studios are rare in many countries, which makes the scenario hard to relate to. Contrast this Item 22, which taps into a similar behavior but is probably more relatable.

21B<sup>†</sup> I would go home and think, "I'll get to know them next time."

Asking a reverse-coded item about the same scenario helps us test whether people are just agreeing with all the items. This can reveal cultural differences in how people answer surveys. However, this item did not pass checks of equivalence across cultures and was removed from the final version.

22\* You're in a math class with 100 people in it. After class on Wednesday, you go the cafeteria to eat lunch. You sit at a table, and you are studying your notes from the class. The cafeteria is crowded, so another person sits down at your table. You recognize the person is in your class, but you've never talked to that person and don't know the person's name. You could continue studying your notes, or you could talk to that person "Hey, how do you like the class?"

**I would start a conversation with that person.**

Responsibilistic cultures prioritize close relationships. Building relationships with weak ties is not usually a priority. This item specifies being in the same large group (the same school and the same class) but also that the other person is not a close tie. This is based on our theorizing that responsibilism is more about tight ties in specific relationships, not being in the same large group.

23 In college, you meet a boy/girl from another country in one of your classes. You like him/her a lot, and you start dating. After you are together for 6 months, you go out to dinner with several of your friends. At dinner, your friends say, "He/she doesn't really like you. He's just going to go back to his country and leave you." They make you so sad that you start to cry after you go back home.

These people aren't really your friends. I should find friends that are more supportive.

This item contrasts warm fuzzies versus long-term duties about people's objective outcomes. If people think that relationships are binding and long lasting, they should expect that the friendship will last through the harsh feedback. If people think that expressing emotional support is a primary function of relationships, they should agree with this item. Collaborators in some cultures said dating is not common there. Thus, we created a similar scenario about having a dream of becoming a writer.

24<sup>†</sup> You're driving in town, and you stop at a red light. After you stop, the car behind you crashes into your car. The accident is serious enough to make it so that you have to repair the car, but no one is hurt. You see the other driver get out of the car. You were driving normally, and the accident was not your fault.

The other driver will come out and try to blame you, saying that you stopped too quickly, so you have to pay for it.

This item taps into people's wariness of strangers, which we theorized would be higher in responsibilistic cultures. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One possible explanation is that people in many cultures do not have experience driving their own car.

25<sup>†</sup> You are a judge, and one day you hear a case and think the person is guilty. But before you decide, your father calls you and says the person is his friend. He asks you to decide "not guilty."

My father would influence my decision.

This item pits a duty to a close family member against a duty to an impersonal system of justice. However, this item did not pass checks of equivalence across cultures and was removed from the final version. One explanation is that few people have experience being a judge. Thus, this item is probably hard to relate to.

26\*

Imagine you've worked for a company—you like your job, and you like your colleagues. But you haven't gotten a raise for 5 years. Now you're thinking about asking your boss for a raise. Of course, it could anger your boss and make working there uncomfortable.

I would still ask my boss for a raise.

This scenario asks about risking disrupting a work relationship. In systems where relationships (even work relationships) tend to be longer-lasting, the consequences of upsetting the relationship will be harsher<sup>20</sup>. In contrast, in societies like the US, where relationships tend to be shorter, the costs of asserting yourself and risking upsetting the relationship are smaller. We adapted this item from a study by Peng and colleagues, who argued that concrete scenarios are better at detecting cultural differences than abstract statements<sup>10</sup>.

Note: This is the full scale before tests of reliability, item-total correlations, and external validity. These tests led to a shortened version of the scale with stronger results on reliability, tests of equivalence across cultures, and external validity.

\*Items marked with asterisks are reverse coded so that higher scores represent more responsibility.

†Crosses mark items that we removed from the final scale because they did not pass tests of reliability, equivalence, or external validity.

**Table S27A. Sources and sample sizes for previous collectivism scales used to test individual-level correlations with responsibilism.**

| Survey Name                                                          | Source                 | <i>N</i>         | Cultures and Platforms Data Collected                                                                        |
|----------------------------------------------------------------------|------------------------|------------------|--------------------------------------------------------------------------------------------------------------|
| Self-Construal Scale                                                 | Singelis, 1994         | 173              | University of Texas Dallas                                                                                   |
| Horizontal & Vertical Individualism & Collectivism                   | Singelis et al., 1995  | 807 <sup>a</sup> | Lebanon 104, Spain 53, China 74, Mexico 68, Chile 107, Argentina 94, South Africa 51, Peru 28, U Chicago 154 |
| Individualism-Collectivism scale                                     | Oyserman, 1993         | 100              | China 74, UK Prolific 26                                                                                     |
| Relational-Interdependent Self-Construal Scale                       | Cross et al., 2000     | 405              | Canada 331, China 74                                                                                         |
| Individualism-Collectivism Scale                                     | Hui, 1988              | 101              | UK Prolific                                                                                                  |
| Individualism-Collectivism Scale                                     | Earley, 1993           | 96               | UK Prolific                                                                                                  |
| Culture and Identity Research Network Self-Construal Scale Version 3 | Vignoles et al., 2016  | 101              | UK Prolific                                                                                                  |
| Self-Construal Scale                                                 | Gudykunst et al., 1996 | 96               | UK Prolific                                                                                                  |

Note: We collected these samples to test the **individual-level** correlations between the responsibilism scale and previous collectivism scales. These are the 8 most-cited collectivism scales from the Oyserman meta-analysis, as well as the more recent scales from Cross and Vignoles. The next tables describe the samples used to test **culture-level** correlations between collectivism and external markers of collectivism. <sup>a</sup>The sample size for the Singelis 1995 scale is larger than the others because a member of the research team often uses this scale and so included it in the sites he collected.

**Table S27B. Sample demographics for testing previous collectivism scales' culture-level predictive power before inclusion criteria (part 1).**

| <b>Culture</b> | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b> | <b>Site</b> | <b>Compensation</b> |
|----------------|----------|---------------|------------|-----------|-----------------|-----------------|-------------|---------------------|
| Algeria        | 156      | Student       | 28.77      | 8.03      | 18.7            | Arabic          | Various     | Payment             |
| Argentina      | 194      | Student       | 30.03      | 10.28     | 41.2            | Spanish         | Various     | Payment             |
| Australia      | 103      | Student       | 28.38      | 9.69      | 54.5            | English         | Various     | Payment             |
| Austria        | 82       | Stud./Adult   | 28.45      | 7.62      | 41.3            | German          | Various     | Payment             |
| Belgium        | 60       | Stud./Adult   | 26.37      | 7.93      | 51.7            | French          | Various     | Payment             |
| Brazil         | 109      | Student       | 21.72      | 6.51      | 77.1            | Portuguese      | Aracaju     | Course Credit       |
| Canada         | 275      | Student       | 24.11      | 5.74      | 62.9            | English         | Various     | Payment             |
| Chile          | 62       | Student       | 25.03      | 3.20      | 22.6            | Spanish         | Various     | Payment             |
| China          | 70       | Student       | 19.14      | 1.60      | 74.3            | Simpl. Chinese  | Wuhan       | Payment             |
| Colombia       | 222      | Student       | 27.76      | 8.63      | 38.7            | Spanish         | Various     | Payment             |
| Czech Rep.     | 61       | Stud./Adult   | 28.24      | 6.32      | 38.3            | Czech           | Various     | Payment             |
| Denmark        | 61       | Stud./Adult   | 32.97      | 11.05     | 39.3            | English         | Various     | Payment             |
| Egypt          | 70       | Stud./Adult   | 23.15      | 6.35      | 92.6            | Arabic          | Menoufia    | Volunteer           |
| Ethiopia       | 165      | Student       | 26.24      | 4.48      | 13.5            | Amharic         | Various     | Payment             |
| Finland        | 62       | Stud./Adult   | 34.82      | 9.89      | 33.9            | Finnish         | Various     | Payment             |
| France         | 97       | Student       | 22.88      | 2.74      | 56.7            | French          | Various     | Payment             |
| Germany        | 117      | Student       | 24.68      | 3.99      | 41.9            | German          | Various     | Payment             |
| Ghana          | 77       | Student       | 21.29      | 4.08      | 72.2            | English         | Accra       | Volunteer           |
| Greece         | 61       | Student       | 26.05      | 5.17      | 34.4            | Greek           | Various     | Payment             |
| India          | 77       | Student       | 19.58      | 1.41      | 59.7            | English         | Various     | Payment             |
| Iran           | 98       | Student       | 24.43      | 7.93      | 54.4            | Persian         | Tehran      | Volunteer           |
| Ireland        | 98       | Stud./Adult   | 32.89      | 11.55     | 69.1            | English         | Various     | Payment             |
| Israel         | 63       | Student       | 25.49      | 4.64      | 43.5            | Hebrew          | Various     | Payment             |
| Italy          | 61       | Student       | 28.48      | 8.55      | 41.0            | Italian         | Various     | Payment             |
| Japan          | 84       | Student       | 20.78      | 3.17      | 63.4            | Japanese        | Various     | Payment             |
| Kenya          | 161      | Student       | 26.42      | 6.52      | 35.2            | English         | Various     | Payment             |

Note: This table presents samples before excluding participants who did not meet the inclusion criteria (such as being born in the local culture) and data-quality checks (such as careless responding). “Various” = internet samples, where participants lived in many places. Participants completed four prior collectivism scales in random order: Singelis et al., 1995; Cross et al., 2000; Earley, 1993; Gudykunst et al., 1996.

**Table S27C. Sample demographics for testing previous collectivism scales' culture-level predictive power before inclusion criteria (part 2).**

| Culture       | N   | Sample      | Age   | SD    | Female % | Language                   | Site         | Compensation |
|---------------|-----|-------------|-------|-------|----------|----------------------------|--------------|--------------|
| Madagascar    | 50  | Student     | 21.88 | 2.96  | 48.0     | French                     | Antananarivo | Payment      |
| Malawi        | 102 | Student     | 26.43 | 6.59  | 52.0     | English                    | Blantyre     | Payment      |
| Mexico        | 61  | Student     | 25.44 | 4.88  | 44.3     | Spanish                    | Various      | Payment      |
| Morocco       | 199 | Student     | 28.35 | 7.93  | 31.1     | French                     | Various      | Payment      |
| Netherlands   | 92  | Student     | 23.40 | 5.44  | 46.7     | Dutch                      | Various      | Payment      |
| New Zealand   | 124 | Stud./Adult | 28.98 | 10.96 | 60.7     | English                    | Various      | Payment      |
| Nigeria       | 321 | Student     | 30.15 | 7.69  | 46.9     | English                    | Various      | Payment      |
| Norway        | 65  | Stud./Adult | 30.02 | 9.01  | 35.9     | English                    | Various      | Payment      |
| Oman          | 87  | Student     | 23.59 | 4.42  | 68.2     | Arabic                     | Various      | Payment      |
| Pakistan      | 64  | Stud./Adult | 31.39 | 9.85  | 12.5     | English                    | Punjab       | Payment      |
| Peru          | 466 | Student     | 26.74 | 8.53  | 41.8     | Spanish                    | Lima         | Payment      |
| Poland        | 59  | Student     | 23.08 | 3.92  | 28.8     | Polish                     | Various      | Payment      |
| Portugal      | 60  | Student     | 24.20 | 3.54  | 41.7     | Portuguese                 | Various      | Payment      |
| Saudi Arabia  | 93  | Student     | 24.79 | 6.91  | 28.9     | Arabic                     | Tabuk        | Volunteer    |
| South Africa  | 60  | Student     | 25.63 | 5.17  | 63.3     | English                    | Various      | Payment      |
| Spain         | 87  | Student     | 26.08 | 6.53  | 45.3     | Spanish                    | Various      | Payment      |
| Sri Lanka     | 250 | Student     | 28.58 | 8.08  | 20.4     | Sinhala                    | Various      | Payment      |
| Sweden        | 96  | Stud./Adult | 32.09 | 10.33 | 30.2     | Swedish                    | Various      | Payment      |
| Switzerland   | 65  | Stud./Adult | 29.91 | 9.96  | 35.4     | German,<br>French, Italian | Various      | Payment      |
| Thailand      | 113 | Student     | 21.17 | 2.23  | 59.3     | Thai                       | Bangkok      | Payment      |
| Turkey        | 243 | Student     | 21.99 | 2.26  | 57.7     | Turkish                    | Istanbul     | Volunteer    |
| UK            | 119 | Student     | 33.62 | 11.20 | 60.2     | English                    | Various      | Payment      |
| United States | 329 | Student     | 32.04 | 11.30 | 47.3     | English                    | Various      | Payment      |
| Venezuela     | 303 | Student     | 34.18 | 12.39 | 40.4     | Spanish                    | Various      | Payment      |

Note: This table presents samples before excluding participants who did not meet the inclusion criteria (such as being a current student and being born in the local culture) and data-quality checks (such as careless responding as measured by low intra-individual response variability). “Various” = internet samples, where participants lived in many places. Participants completed four prior collectivism scales in random order: Singelis et al., 1995; Cross et al., 2000; Earley, 1993; Gudykunst et al., 1996.

**Table S27D. Sample demographics for testing previous collectivism scales' culture-level predictive power after inclusion criteria (part 1).**

| <b>Culture</b> | <b>N</b> | <b>Sample</b> | <b>Age</b> | <b>SD</b> | <b>Female %</b> | <b>Language</b> | <b>Site</b> | <b>Compensation</b> |
|----------------|----------|---------------|------------|-----------|-----------------|-----------------|-------------|---------------------|
| Algeria        | 51       | Student       | 24.31      | 5.93      | 13.7            | Arabic          | Various     | Payment             |
| Argentina      | 60       | Student       | 24.87      | 7.45      | 45.0            | Spanish         | Various     | Payment             |
| Australia      | 57       | Student       | 29.32      | 10.05     | 64.9            | English         | Various     | Payment             |
| Austria        | 50       | Stud./Adult   | 27.30      | 6.14      | 34.0            | German          | Various     | Payment             |
| Belgium        | 48       | Stud./Adult   | 26.60      | 8.11      | 52.1            | French          | Various     | Payment             |
| Brazil         | 106      | Student       | 21.60      | 6.46      | 76.4            | Portuguese      | Aracaju     | Course Credit       |
| Canada         | 109      | Student       | 25.79      | 6.83      | 65.1            | English         | Various     | Payment             |
| Chile          | 52       | Student       | 24.94      | 2.81      | 21.2            | Spanish         | Various     | Payment             |
| China          | 54       | Student       | 19.04      | 1.48      | 77.8            | Simpl. Chinese  | Wuhan       | Payment             |
| Colombia       | 103      | Student       | 23.28      | 5.65      | 38.8            | Spanish         | Various     | Payment             |
| Czech Rep.     | 48       | Stud./Adult   | 27.89      | 6.22      | 37.5            | Czech           | Various     | Payment             |
| Denmark        | 45       | Stud./Adult   | 33.31      | 11.15     | 37.8            | English         | Various     | Payment             |
| Egypt          | 56       | Stud./Adult   | 22.51      | 4.61      | 96.4            | Arabic          | Menoufia    | Volunteer           |
| Ethiopia       | 58       | Student       | 23.78      | 2.24      | 10.3            | Amharic         | Various     | Payment             |
| Finland        | 50       | Stud./Adult   | 35.68      | 10.00     | 32.0            | Finnish         | Various     | Payment             |
| France         | 50       | Student       | 22.56      | 2.57      | 60.0            | French          | Various     | Payment             |
| Germany        | 69       | Student       | 25.25      | 3.97      | 37.7            | German          | Various     | Payment             |
| Ghana          | 53       | Student       | 21.78      | 4.57      | 69.2            | English         | Accra       | Volunteer           |
| Greece         | 47       | Student       | 26.02      | 5.20      | 40.4            | Greek           | Various     | Payment             |
| India          | 47       | Student       | 19.57      | 1.54      | 72.3            | English         | Various     | Payment             |
| Iran           | 70       | Student       | 23.41      | 7.53      | 62.9            | Persian         | Tehran      | Volunteer           |
| Ireland        | 63       | Stud./Adult   | 34.98      | 11.56     | 68.3            | English         | Various     | Payment             |
| Israel         | 49       | Student       | 25.05      | 4.62      | 46.9            | Hebrew          | Various     | Payment             |
| Italy          | 44       | Student       | 28.39      | 8.46      | 31.8            | Italian         | Various     | Payment             |
| Japan          | 68       | Student       | 20.41      | 1.20      | 67.6            | Japanese        | Various     | Payment             |
| Kenya          | 51       | Student       | 24.18      | 3.68      | 33.3            | English         | Various     | Payment             |

Note: This table presents the demographics after exclusion criteria (such as being a current student, being born in the local culture, and having grown up in the local culture) and data-quality checks (such as careless responding as measured by low intra-individual response variability). “Various” = internet samples, where participants lived in many places.

**Table S27E. Sample demographics for testing previous collectivism scales' culture-level predictive power after inclusion criteria (part 2).**

| Culture       | N   | Sample      | Age   | SD    | Female % | Language                   | Site         | Compensation |
|---------------|-----|-------------|-------|-------|----------|----------------------------|--------------|--------------|
| Madagascar    | 46  | Student     | 21.50 | 2.46  | 45.7     | French                     | Antananarivo | Payment      |
| Malawi        | 102 | Student     | 26.43 | 6.59  | 52.0     | English                    | Blantyre     | Payment      |
| Mexico        | 51  | Student     | 24.45 | 3.72  | 49.0     | Spanish                    | Various      | Payment      |
| Morocco       | 57  | Student     | 23.02 | 3.43  | 35.1     | French                     | Various      | Payment      |
| Netherlands   | 63  | Student     | 23.17 | 5.93  | 54.0     | Dutch                      | Various      | Payment      |
| New Zealand   | 61  | Stud./Adult | 31.89 | 12.85 | 62.3     | English                    | Various      | Payment      |
| Nigeria       | 81  | Student     | 25.47 | 6.08  | 35.0     | English                    | Various      | Payment      |
| Norway        | 50  | Stud./Adult | 30.14 | 9.20  | 38.0     | English                    | Various      | Payment      |
| Oman          | 60  | Student     | 22.87 | 3.47  | 71.7     | Arabic                     | Various      | Payment      |
| Pakistan      | 64  | Stud./Adult | 31.39 | 9.85  | 12.5     | English                    | Punjab       | Payment      |
| Peru          | 164 | Student     | 23.12 | 6.54  | 42.7     | Spanish                    | Lima         | Payment      |
| Poland        | 43  | Student     | 22.98 | 4.33  | 34.9     | Polish                     | Various      | Payment      |
| Portugal      | 43  | Student     | 23.91 | 3.29  | 44.2     | Portuguese                 | Various      | Payment      |
| Saudi Arabia  | 50  | Student     | 21.04 | 1.16  | 0.00     | Arabic                     | Tabuk        | Volunteer    |
| South Africa  | 55  | Student     | 25.35 | 4.77  | 65.5     | English                    | Various      | Payment      |
| Spain         | 58  | Student     | 25.64 | 6.29  | 44.8     | Spanish                    | Various      | Payment      |
| Sri Lanka     | 70  | Student     | 24.00 | 3.95  | 22.9     | Sinhala                    | Various      | Payment      |
| Sweden        | 57  | Stud./Adult | 33.57 | 10.44 | 29.8     | Swedish                    | Various      | Payment      |
| Switzerland   | 41  | Stud./Adult | 29.29 | 9.74  | 36.6     | German,<br>French, Italian | Various      | Payment      |
| Thailand      | 97  | Student     | 20.89 | 1.56  | 60.8     | Thai                       | Bangkok      | Payment      |
| Turkey        | 194 | Student     | 22.01 | 2.40  | 60.3     | Turkish                    | Istanbul     | Volunteer    |
| UK            | 67  | Student     | 35.40 | 11.83 | 61.2     | English                    | Various      | Payment      |
| United States | 105 | Student     | 30.74 | 9.90  | 41.0     | English                    | Various      | Payment      |
| Venezuela     | 93  | Student     | 24.81 | 8.17  | 36.6     | Spanish                    | Various      | Payment      |

Note: This table presents the demographics after exclusion criteria (such as being a current student, being born in the local culture, and having grown up in the local culture) and data-quality checks (such as careless responding as measured by low intra-individual response variability). “Various” = internet samples, where participants lived in many places. Participants completed four prior collectivism scales in random order: Singelis et al., 1995; Cross et al., 2000; Earley, 1993; Gudykunst et al., 1996.



**Table S28. Responsibilism differences by age, gender, and social class.**

|              | <i>B</i> | 95% CI         | <i>SE</i> | $\beta$ | <i>t</i> | <i>P</i> |
|--------------|----------|----------------|-----------|---------|----------|----------|
| Female       | -0.03    | -0.06, -0.01   | 0.01      | -0.03   | -2.68    | 0.007    |
| Age          | -0.01    | -0.006, -0.003 | 0.00      | -0.06   | -6.20    | < 0.001  |
| Social Class | -0.04    | -0.05, -0.03   | 0.00      | -0.11   | -11.83   | < 0.001  |

Note: This table reports a simple regression predicting responsibilism scores from demographics. The sample size is 11,298, which is smaller than the overall sample because some cultures did not have complete demographic data. Female is coded as 0 = male, 1 = female. Age is in years. Participants reported social class using the ladder scale from 0 (*bottom*) to 10 (*top*) compared to people in their country. We saved unstandardized residuals from this analysis to estimate overall culture scores in responsibilism after accounting for demographic differences. Cultural differences between non-Western and Western cultures were significant before,  $t(65) = 16.30$ ,  $P < 0.001$ ,  $d = 4.58$ , 95% CI [3.61, 5.53], and after adjusting for demographics  $t(65) = 16.41$ ,  $P < 0.001$ ,  $d = 4.61$ , 95% CI [3.64, 5.57].

**Table S29. Cultural differences are similar before (top) and after (bottom) controlling for demographics.**

| Model                        | Variable                       | <i>B</i> | <i>SE</i> | $\beta$ | <i>t</i> | <i>P</i> |
|------------------------------|--------------------------------|----------|-----------|---------|----------|----------|
| Basic Model                  | West vs. Non-West <sup>a</sup> | 0.42     | 0.01      | 0.54    | 59.43    | < 0.001  |
| < 24 Years Old               | West vs. Non-West <sup>a</sup> | 0.44     | 0.01      | 0.56    | 50.85    | < 0.001  |
| Controlling for Demographics | Female                         | 0.00     | 0.01      | 0.00    | -0.15    | 0.879    |
|                              | Age                            | 0.00     | 0.00      | -0.04   | -3.93    | < 0.001  |
|                              | Social Class                   | -0.02    | 0.00      | -0.06   | -6.40    | < 0.001  |
|                              | West vs. Non-West <sup>a</sup> | 0.42     | 0.01      | 0.53    | 58.25    | < 0.001  |

Note: <sup>a</sup>West vs. Non-West is coded as 0 = Western culture, 2 = non-Western culture. [Supplementary Section 24](#) defines Western and non-Western cultures in detail. This table reports a simple regression measuring West vs. non-West cultural differences in a basic model not taking into account demographic differences (top) and in a model that takes into account demographic differences (bottom). The size of cultural differences is similar before and after taking into account demographic differences, which suggests that cultural differences are independent of demographic differences between samples in different cultures. The sample size of the top and bottom regressions is 8,673, which is smaller than the overall sample because (a) some cultures did not have complete demographic data and (b) this analysis compares West and non-Western cultures, which does not include “sometimes Western” cultures. An analysis in [Supplementary Section 3](#) includes “sometimes Western” cultures and reaches a similar conclusion. The sample size in the middle regression (limited to < 24 years old) is 5,782. Female is coded as 0 = male, 1 = female. Age is in years. Participants reported social class using the ladder scale from 0 (*bottom*) to 10 (*top*) compared to people in their country.

**Table S30. Spatial auto-correlation: Responsibilism predicts behavioral differences that tap into tight responsibilities in close relationships after controlling for latitude and longitude.**

| Marker                                                                                      | $r_{\text{partial}}$ | $P_{\text{partial}}$ | $P_{\text{partial}}$<br>Adjusted <sup>a</sup> | 95% CI      | Cultures |
|---------------------------------------------------------------------------------------------|----------------------|----------------------|-----------------------------------------------|-------------|----------|
| <b>Tight Ties in Close Relationships</b>                                                    |                      |                      |                                               |             |          |
| Global Collectivism Index                                                                   | 0.71                 | < 0.001              | < 0.001                                       | 0.54 0.81   | 99       |
| Kinship Intensity Index                                                                     | 0.62                 | < 0.001              | < 0.001                                       | 0.50 0.72   | 99       |
| Blood Donations Family                                                                      | 0.31                 | 0.003                | 0.004                                         | 0.16 0.44   | 91       |
| Lying for Friend in Court                                                                   | 0.61                 | < 0.001              | < 0.001                                       | 0.32 0.77   | 40       |
| <b>Weak Responsibilities to Strangers and Anonymous Systems</b>                             |                      |                      |                                               |             |          |
| UN Diplomat Parking Tickets                                                                 | 0.56                 | < 0.001              | < 0.001                                       | 0.33 0.69   | 89       |
| Dishonest Claims in Private Dice Game                                                       | 0.77                 | < 0.001              | < 0.001                                       | 0.50 0.93   | 19       |
| Societal Discrimination Against Outgroups                                                   | 0.81                 | < 0.001              | < 0.001                                       | 0.68 0.88   | 96       |
| <b>Pro-Sociality with Strangers</b>                                                         |                      |                      |                                               |             |          |
| Blood Donations Strangers                                                                   | -0.72                | < 0.001              | < 0.001                                       | -0.80 -0.59 | 89       |
| Wallet Return Rate to Strangers                                                             | -0.78                | < 0.001              | < 0.001                                       | -0.89 -0.57 | 39       |
| Contributions in Anon. Economic Game                                                        | -0.45                | 0.121                | 0.132                                         | -0.88 0.26  | 15       |
| Contributions in Anon. Econ. Game (1 <sup>st</sup> Round)                                   | -0.37                | 0.218                | 0.218                                         | -0.77 0.51  | 15       |
| <b>Basing Life Satisfaction on Internal Feelings (as Opposed to Social Roles or Status)</b> |                      |                      |                                               |             |          |
| Life Satisfaction Correlation w/ Emotion                                                    | -0.40                | 0.021                | 0.025                                         | -0.71 -0.06 | 35       |

Note: This table reports the partial correlations between culture-level responsibilism scores and markers of external validity after controlling for longitude and latitude. These measure responsibilities in close relationships (such as lying for a friend in court) versus favorable treatment of strangers (such as returning a wallet to a stranger). Controlling for latitude and longitude helps account for the possibility that two variables might be significantly correlated only because they tend to be in the same physical locations (spatial auto-correlation). If a correlation remains significant after controlling for latitude and longitude, it suggests that the relationship is robust beyond simply appearing together in physical space. Correlations in bold are statistically significant. Extended Data Tables 1 and 2 outline the data sources and theoretical rationales. <sup>a</sup>This column reports  $P$  values adjusted for multiple comparisons using the Benjamini-Hochberg method.

**Table S31. Spatial auto-correlation: Culture-level partial correlations of responsibilism with ecological and historical conditions after controlling for latitude and longitude.**

| Characteristic                | $r_{\text{partial}}$ | $P_{\text{partial}}$ | $P_{\text{partial}}$<br>Adjusted <sup>a</sup> | 95% CI |       | Cultures |
|-------------------------------|----------------------|----------------------|-----------------------------------------------|--------|-------|----------|
| Development and Health        |                      |                      |                                               |        |       |          |
| GDP per Capita 1918 Log       | <b>-0.74</b>         | < 0.001              | < 0.001                                       | -0.86  | -0.40 | 33       |
| GDP per Capita 1950 Log       | <b>-0.58</b>         | < 0.001              | < 0.001                                       | -0.82  | -0.18 | 79       |
| GDP per Capita 1970 Log       | <b>-0.59</b>         | < 0.001              | < 0.001                                       | -0.78  | -0.28 | 79       |
| GDP per Capita 1990 Log       | <b>-0.63</b>         | < 0.001              | < 0.001                                       | -0.75  | -0.43 | 91       |
| GNI per Capita 2019 Log       | <b>-0.56</b>         | < 0.001              | < 0.001                                       | -0.71  | -0.33 | 98       |
| % Urban Population 2013       | <b>-0.41</b>         | < 0.001              | < 0.001                                       | -0.59  | -0.15 | 98       |
| Historical Disease Prevalence | <b>0.61</b>          | < 0.001              | < 0.001                                       | 0.38   | 0.74  | 100      |
| Religion                      |                      |                      |                                               |        |       |          |
| % Catholic 2000               | <b>-0.27</b>         | 0.009                | 0.012                                         | -0.45  | -0.11 | 95       |
| % Protestant 2000             | <b>-0.38</b>         | < 0.001              | < 0.001                                       | -0.55  | -0.13 | 95       |
| % Muslim 2000                 | <b>0.48</b>          | < 0.001              | < 0.001                                       | 0.35   | 0.60  | 95       |
| Western Church Exposure       | <b>-0.71</b>         | < 0.001              | < 0.001                                       | -0.81  | -0.50 | 94       |
| Eastern Church Exposure       | -0.13                | 0.204                | 0.216                                         | -0.26  | 0.01  | 94       |
| Ecology                       |                      |                      |                                               |        |       |          |
| Population Density in 1CE     | -0.09                | 0.430                | 0.430                                         | -0.26  | 0.10  | 82       |
| Population Density in 1500CE  | <b>-0.21</b>         | 0.050                | 0.060                                         | -0.40  | -0.02 | 95       |
| % Rice Harvested Area 2000    | <b>0.23</b>          | 0.026                | 0.033                                         | -0.05  | 0.41  | 100      |
| Historical Herding            | <b>-0.27</b>         | 0.009                | 0.012                                         | -0.45  | -0.04 | 98       |
| Climatic Demands              | <b>-0.36</b>         | < 0.001              | < 0.001                                       | -0.50  | -0.16 | 99       |
| Historical War                | 0.21                 | 0.082                | 0.092                                         | -0.02  | 0.40  | 74       |

Note: This table tests potential social and ecological causes of responsibilism after controlling for latitude and longitude (using partial correlations). Controlling for latitude and longitude helps account for the possibility that two variables might be significantly correlated only because they tend to be in the same physical locations (spatial auto-correlation). If a correlation remains significant after controlling for latitude and longitude, it suggests that the relationship is robust beyond simply appearing together in physical space. Correlations in bold are statistically significant. Extended Data Table 4 outlines the data sources and theoretical rationales. <sup>a</sup>In this column, we report  $P$  values that are adjusted for multiple comparisons using the Benjamini-Hochberg method, which makes false positives less likely.

**Table S32. Wordings for the five most-cited collectivism scales.**

| <b>Singelis, 1994, Interdependence</b>                                                                        |
|---------------------------------------------------------------------------------------------------------------|
| 1. I have respect for the authority figures with whom I interact.                                             |
| 2. It is important for me to maintain harmony within my group.                                                |
| 3. My happiness depends on the happiness of those around me.                                                  |
| 4. I would offer my seat in a bus to my professor.                                                            |
| 5. I respect people who are modest about themselves.                                                          |
| 6. I will sacrifice my self-interest for the benefit of the group I am in.                                    |
| 7. I often have the feeling that my relationships with others are more important than my own accomplishments. |
| 8. I should take into consideration my parents' advice when making education/career plans.                    |
| 9. It is important to me to respect decisions made by the group.                                              |
| 10. I will stay in a group if they need me, even when I'm not happy with the group.                           |
| 11. If my brother or sister fails, I feel responsible.                                                        |
| 12. Even when I strongly disagree with group members, I avoid argument.                                       |
| <b>Singelis et al., 1995, Vertical-Horizontal Collectivism</b>                                                |
| 1. I hate to disagree with others in my group. (Vertical)                                                     |
| 2. Children should feel honored if their parents receive a distinguished award. (Vertical)                    |
| 3. My happiness depends very much on the happiness of those around me. (Horizontal)                           |
| 4. If a relative were in financial difficulty, I would help within my means. (Horizontal)                     |
| 5. The well-being of my coworkers is important to me. (Horizontal)                                            |
| 6. We should keep our aging parents with us at home. (Vertical)                                               |
| 7. I would sacrifice an activity I enjoy very much if my family did not approve of it. (Vertical)             |
| 8. Before taking a major trip, I consult with most members of my family and many friends. (Vertical)          |
| 9. I usually sacrifice my self-interest for the benefit of the group. (Vertical)                              |
| 10. I feel good when I cooperate with others. (Horizontal)                                                    |
| 11. I would do what would please my family, even if I detested that activity. (Vertical)                      |
| 12. It is important to maintain harmony within my group. (Horizontal)                                         |
| 13. Children should be taught to place duty before pleasure. (Vertical)                                       |
| 14. I like sharing little things with my neighbors. (Horizontal)                                              |
| 15. If a coworker gets a prize, I would feel proud. (Horizontal)                                              |
| 16. To me, pleasure is spending time with others. (Horizontal)                                                |

---

**Gudykunst et al., 1996, Interdependence Scale**

---

1. I consult with others before making important decisions.
  2. I consult with co-workers on work-related matters.
  3. I will sacrifice my self-interest for the benefit of my group.
  4. I stick with my group even through difficulties.
  5. I respect decisions made by my group.
  6. I will stay in a group if they need me, even when I am not happy with the group.
  7. I maintain harmony in the groups of which I am a member.
  8. I respect the majority's wishes in groups of which I am a member.
  9. I remain in the groups of which I am a member if they need me, even though I am dissatisfied with them.
  10. I try to abide by customs and conventions at work.
  11. I give special consideration to others' personal situations so I can be efficient at work.
  12. It is better to consult with others and get their opinions before doing anything.
  13. It is important to consult close friends and get their ideas before making a decision.
  14. My relationships with others are more important than my accomplishments.
- 

**Cross et al., 2000, Relational-Interdependent Self-Construal Scale**

---

1. My close relationships are an important reflection of who I am.
  2. When I feel very close to someone, it often feels to me like that person is an important part of who I am.
  3. Overall, my close relationships have very little to do with how I feel about myself.  
(Reversed)
  4. I think one of the most important parts of who I am can be captured by looking at my close friends and understanding who they are.
  5. When I think of myself, I often think of my close friends or family also.
  6. When I establish a close friendship with someone, I usually develop a strong sense of identification with that person.
  7. If a person hurts someone close to me, I feel hurt as well.
  8. My close relationships are unimportant to my sense of what kind of person I am. (Reversed)
  9. My sense of pride comes from knowing who I have as close friends.
  10. In general, my close relationships are an important part of my self-image.
  11. I usually feel a strong sense of pride when someone close to me has an important accomplishment.
-

---

### Earley, 1993, Collectivism Scale

---

1. Employees like to work in a group rather than by themselves.
  2. If a group is slowing me down, it is better to leave it and work alone. (Reversed)
  3. To be superior, a person must stand alone. (Reversed)
  4. One does better work working alone than in a group. (Reversed)
  5. I would rather struggle through a personal problem by myself than discuss it with my friends. (Reversed)
  6. An employee should accept the group's decision even when personally he or she has a different opinion.
  7. Problem solving by groups gives better results than problem solving by individuals.
  8. The needs of people close to me should take priority over my personal needs.
- 

Note: These are the five most-cited collectivism scales among the scales included in an earlier meta-analysis<sup>11</sup>. [Table S9](#) lists the citation counts. The Gudykunst scale and the two Singelis scales also have an individualism sub-scale (not listed here). We did not include the individualism sub-scales because (a) there is debate about whether individualism is a part of collectivism or separate from it and (b) including those items would have forced us to cut entire other scales (such as the Earley scale) from the data collection in order to avoid overburdening participants with too many items.

**Table S33. Skewness and kurtosis of culture-level variables.**

| <b>Culture-Level Variables</b>                       | <b>Skewness</b> | <b>Kurtosis</b> | <b>N</b> |
|------------------------------------------------------|-----------------|-----------------|----------|
| Responsibilism                                       | -0.60           | -0.36           | 100      |
| Responsibilism Item 1                                | -0.001          | -0.82           | 100      |
| Responsibilism Item 2                                | -0.18           | -0.10           | 100      |
| Responsibilism Item 3                                | 0.11            | -0.17           | 100      |
| Responsibilism Item 4                                | -0.21           | -0.68           | 100      |
| Responsibilism Item 5                                | -0.53           | -0.25           | 100      |
| Responsibilism Item 6                                | 0.09            | -1.15           | 100      |
| Responsibilism Item 7                                | 0.17            | 0.11            | 100      |
| Responsibilism Item 8                                | -1.09           | 1.11            | 100      |
| Responsibilism Item 9                                | 0.04            | -0.46           | 100      |
| Responsibilism Item 10                               | 0.25            | -0.74           | 100      |
| Responsibilism Item 11                               | 0.45            | -0.35           | 100      |
| Responsibilism Item 12                               | 0.55            | 0.15            | 100      |
| Responsibilism Item 13                               | -0.13           | -0.08           | 100      |
| Responsibilism Item 14                               | 0.52            | 0.48            | 100      |
| Responsibilism (Five Item)                           | -0.50           | -0.68           | 100      |
| Family Responsibilities                              | -0.22           | -0.71           | 100      |
| Guarding Personal Details                            | -0.20           | -0.39           | 100      |
| Tight Friendship                                     | 0.12            | -0.12           | 100      |
| Stranger Danger                                      | -0.09           | -0.13           | 100      |
| Fitting In                                           | 0.10            | -0.25           | 100      |
| Horizontal Collectivism (Singelis et al., 1995)      | -1.74           | 4.58            | 50       |
| Vertical Collectivism (Singelis et al., 1995)        | 0.73            | -0.63           | 50       |
| Collectivism (Vert. + Horiz.; Singelis et al., 1995) | 0.34            | 0.36            | 50       |
| Interdependence (Gudykunst et al., 1996)             | -0.53           | -0.51           | 50       |
| Collectivism (Earley, 1993)                          | -0.68           | 1.23            | 50       |
| Relational-Int. Self-Construal (Cross et al., 2000)  | -0.47           | -0.02           | 50       |
| Interdependence (Singelis, 1994)                     | -0.57           | -0.36           | 35       |
| Independence (Singelis, 1994)                        | 0.44            | -0.04           | 35       |
| Interdependence – Independence (Singelis, 1994)      | -0.70           | 0.65            | 35       |

Note: We compared skewness to a suggested cutoff of 2 and kurtosis to a cutoff of  $7^{106}$ . All of the variables in this table were under the cutoffs.



**Table S34A. Skewness and kurtosis of culture-level external indicators (part 1).**

| <b>Culture-Level Indicators</b>                           | <b>Skewness</b> | <b>Kurtosis</b> | <b>N</b> |
|-----------------------------------------------------------|-----------------|-----------------|----------|
| Global Collectivism Index                                 | 0.14            | -1.03           | 99       |
| Kinship Intensity Index                                   | 0.17            | -1.50           | 99       |
| Blood Donations Family                                    | 1.18            | -0.07           | 91       |
| Lying for Friend in Court                                 | 0.65            | -0.42           | 40       |
| UN Diplomat Parking Tickets                               | 0.19            | -0.83           | 89       |
| Dishonest Claims in Private Dice Game                     | -0.29           | -0.79           | 19       |
| Societal Discrimination Against Outgroups                 | -0.59           | -0.56           | 96       |
| Blood Donations Strangers                                 | 0.45            | -1.20           | 89       |
| Wallet Return Rate to Strangers                           | -0.15           | -1.17           | 39       |
| Contributions in Anon. Economic Game                      | -0.48           | -0.73           | 15       |
| Contributions in Anon. Econ. Game (1 <sup>st</sup> Round) | 0.32            | -1.11           | 15       |
| Life Satisfaction Correlation with Emotion                | -1.04           | 0.97            | 35       |
| Democracy Index 2021                                      | -0.39           | -0.88           | 95       |
| Corruption Index 2019                                     | -0.48           | -0.96           | 97       |
| Global Innovation Index                                   | 0.29            | -0.99           | 86       |
| Economic Freedom Index 2021 (Neoliberalism)               | -0.43           | 1.19            | 95       |
| Holistic Thought                                          | 0.22            | -0.25           | 26       |
| Migrant Integration Policy Index                          | 0.05            | 0.53            | 48       |
| Covid-19 Deaths per Capita                                | -0.75           | -0.18           | 94       |
| Death Rate from Drug Use Disorders                        | 4.87            | 31.38           | 98       |
| GDP per Capita 1918 Log                                   | -0.07           | -1.17           | 33       |
| GDP per Capita 1950 Log                                   | 0.34            | -0.29           | 79       |
| GDP per Capita 1970 Log                                   | -0.10           | -1.04           | 79       |
| GDP per Capita 1990 Log                                   | -0.39           | -1.04           | 91       |
| 2019 Log GNI per Capita                                   | -0.67           | -0.29           | 98       |
| % Urban Population 2013                                   | -0.41           | -0.78           | 98       |
| Historical Disease Prevalence                             | -0.12           | -1.16           | 100      |
| % Catholic 2000                                           | 0.96            | -0.69           | 95       |
| % Protestant 2000                                         | 1.99            | 3.59            | 95       |
| % Muslim 2000                                             | 1.65            | 1.13            | 95       |
| Western Church Exposure                                   | 1.16            | 0.05            | 94       |
| Eastern Church Exposure                                   | 4.31            | 19.63           | 94       |

Note: We compared skewness to a suggested cutoff of 2 and kurtosis to a cutoff of 7<sup>106</sup>. Eastern Church exposure and death rates from drug use disorders exceeded these cutoffs, so we ran additional analyses after log transforming these two variables. [Supplemental Section 38](#) reports the results.

**Table S34B. Skewness and kurtosis of culture-level external indicators (part 2).**

| <b>Culture-Level Indicators</b> | <b>Skewness</b> | <b>Kurtosis</b> | <b><i>N</i></b> |
|---------------------------------|-----------------|-----------------|-----------------|
| Absolute Latitude               | 0.02            | -1.15           | 97              |
| Longitude                       | -0.07           | -0.03           | 97              |
| Population Density in 1CE       | 2.19            | 6.89            | 82              |
| Population Density in 1500CE    | 1.66            | 2.41            | 95              |
| % Rice Harvested Area 2000      | 1.89            | 2.45            | 100             |
| Historical Herding              | -1.16           | 0.10            | 98              |
| Climatic Demands                | 0.37            | -0.58           | 99              |
| Historical War                  | 1.05            | 0.06            | 74              |
| Relational Mobility             | -0.55           | -0.78           | 55              |
| Norm Tightness                  | 0.65            | 0.36            | 51              |
| Zero-Sum Belief                 | 0.59            | 1.01            | 32              |
| Divorce is Justifiable          | 0.004           | -0.85           | 70              |
| Most People Can be Trusted      | 0.97            | 0.32            | 70              |
| Trust in Strangers              | 0.43            | 0.04            | 70              |
| Uncertainty Avoidance           | -0.24           | -0.80           | 83              |
| Long-Term Orientation           | 0.30            | -0.77           | 73              |
| Power Distance                  | -0.40           | -0.47           | 79              |
| Masculinity                     | -0.05           | 0.33            | 79              |
| Indulgence                      | 0.28            | -0.52           | 70              |
| Gini Index 2013                 | 0.72            | 0.42            | 89              |
| Life Expectancy 2020            | -0.71           | 0.05            | 96              |

Note: We compared skewness to a suggested cutoff of 2 and kurtosis to a cutoff of 7<sup>106</sup>. We did not test for skewness and kurtosis for the variable of whether a language has a future tense marker because it is binary variable. Population density in 1 CE exceeded these cutoffs, so we ran an additional analysis after log transforming the variable. [Supplemental Section 38](#) reports the results.

**Table S35A. Skewness and kurtosis of individual-level variables (part 1).**

| <b>Individual-Level Variables</b>                  | <b>Skewness</b> | <b>Kurtosis</b> | <b>N</b> |
|----------------------------------------------------|-----------------|-----------------|----------|
| Responsibilism                                     | -0.18           | 0.01            | 12,247   |
| Responsibilism (Five Item)                         | -0.21           | -0.21           | 12,247   |
| Family Responsibilities                            | -0.33           | -0.26           | 12,247   |
| Guarding Personal Details                          | -0.17           | -0.44           | 12,247   |
| Tight Friendship                                   | 0.04            | -0.34           | 12,247   |
| Stranger Danger                                    | 0.21            | -0.18           | 12,247   |
| Fitting In                                         | 0.31            | -0.004          | 12,247   |
| Responsibilism (Divergent Validity Sub-Sample)     | -0.17           | -0.15           | 3,450    |
| Interdependence (Singelis, 1994)                   | -0.01           | 0.72            | 173      |
| Independence (Singelis, 1994)                      | 0.19            | 0.15            | 173      |
| Inter. – Independence (Singelis, 1994)             | 0.03            | 0.35            | 173      |
| Horizontal Collectivism (Singelis et al., 1995)    | -0.50           | 0.53            | 890      |
| Vertical Collectivism (Singelis et al., 1995)      | 0.02            | -0.27           | 916      |
| Collectivism (Hor. + Vert.; Singelis et al., 1995) | -0.15           | 0.17            | 916      |
| Horizontal Individualism (Singelis et al., 1995)   | -0.31           | 0.31            | 890      |
| Vertical Individualism (Singelis et al., 1995)     | -0.14           | -0.26           | 890      |
| Collect. – Individualism (Singelis et al., 1995)   | -0.01           | 0.06            | 890      |
| Collectivism (Oyserman, 1993)                      | -0.43           | 0.89            | 108      |
| Individualism (Oyserman, 1993)                     | 0.13            | -0.20           | 108      |
| Coll. – Individualism (Oyserman, 1993)             | -1.13           | 2.95            | 108      |
| Relational Self-Construal (Cross et al., 2000)     | 0.22            | 0.19            | 415      |
| Collectivism (Hui, 1988)                           | -0.28           | 0.60            | 101      |
| Collectivism (Earley, 1993)                        | -0.02           | -0.31           | 96       |
| Interdependence (Gudykunst et al., 1996)           | -0.88           | 1.72            | 96       |
| Independence (Gudykunst et al., 1996)              | -0.39           | 0.22            | 96       |
| Inter. – Independence (Gudykunst et al., 1996)     | -0.20           | 0.91            | 96       |

Note: We compared skewness to a suggested cutoff of 2 and kurtosis to a cutoff of 7<sup>106</sup>. All of the variables in this table were under the cutoffs, so we did not transform the data.

**Table S35B. Skewness and kurtosis of individual-level variables (part 2).**

| <b>Individual-Level Variables</b>              | <b>Skewness</b> | <b>Kurtosis</b> | <b>N</b> |
|------------------------------------------------|-----------------|-----------------|----------|
| Interdependence (Vignoles et al., 2016)        | -0.42           | 0.60            | 126      |
| Difference (Vignoles et al., 2016)             | -0.06           | -0.46           | 126      |
| Self-Containment (Vignoles et al., 2016)       | 0.71            | 0.62            | 126      |
| Self-Direction (Vignoles et al., 2016)         | -0.28           | 0.20            | 126      |
| Self-Reliance (Vignoles et al., 2016)          | -0.19           | -0.16           | 126      |
| Self-Expression (Vignoles et al., 2016)        | 0.04            | -0.34           | 126      |
| Self-Interest (Vignoles et al., 2016)          | -0.20           | -0.13           | 126      |
| Consistency (Vignoles et al., 2016)            | -0.27           | -0.79           | 126      |
| De-Contextualized Self (Vignoles et al., 2016) | -0.36           | -0.04           | 126      |
| Extraversion                                   | 0.14            | -0.83           | 173      |
| Neuroticism                                    | -0.18           | -0.69           | 173      |
| Agreeableness                                  | 0.01            | -0.66           | 173      |
| Conscientiousness                              | -0.28           | -0.27           | 173      |
| Openness                                       | -0.37           | -0.55           | 173      |
| Self-Report Responsibilism Hong Kong           | 0.04            | 0.73            | 516      |
| “Culture-Report” Responsibilism Hong Kong      | 0.23            | 0.04            | 516      |
| Self-Report Responsibilism US                  | -0.16           | 0.15            | 147      |
| “Culture-Report” Responsibilism US             | 0.11            | -0.21           | 147      |
| Responsibilism in China before Pandemic        | -0.44           | 0.38            | 104      |
| Responsibilism in China after Pandemic         | -0.21           | -0.02           | 104      |
| Responsibilism in India before Pandemic        | 0.54            | 0.54            | 28       |
| Responsibilism in India after Pandemic         | 1.03            | 3.57            | 28       |

Note: We compared skewness to a suggested cutoff of 2 and kurtosis to a cutoff of 7<sup>106</sup>. All of the variables in this table were under the cutoffs, so we did not transform any of the variables.

**Table S36A. Testing for equality of variance using Levene's test (part 1).**

|                                                                   | <b>Comparisons</b>                                                                                        | <b><i>F</i></b> | <b><i>P</i></b> | <b><i>N</i></b> |
|-------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|-----------------|-----------------|-----------------|
| Culture-Level<br>Responsibilism                                   | East Asian vs. Western Cultures                                                                           | 1.17            | 0.290           | 24              |
|                                                                   | Non-Western Cultures vs. Western Cultures                                                                 | 0.01            | 0.939           | 68              |
|                                                                   | "Sometimes Western" Cultures vs. Non-Western Cultures                                                     | 0.07            | 0.787           | 83              |
|                                                                   | "Sometimes Western" Cultures vs. Western Cultures                                                         | 0.01            | 0.905           | 49              |
|                                                                   | 18 Former Soviet Bloc and Southern European Cultures vs. 13 Western European Cultures                     | 0.14            | 0.714           | 31              |
| Culture-Level<br>Responsibilism<br>before any Scale<br>Refinement | Non-Western Cultures vs. Western Cultures                                                                 | 0.10            | 0.751           | 68              |
|                                                                   | East Asian vs. Western Cultures                                                                           | 0.003           | 0.956           | 24              |
| Within the US                                                     | US South vs. US North                                                                                     | 0.08            | 0.784           | 576             |
|                                                                   | European Americans vs. Americans with Asian Heritage                                                      | 0.001           | 0.971           | 406             |
|                                                                   | European Americans vs. Americans with Latin American Heritage                                             | 0.43            | 0.514           | 338             |
| Within South Africa                                               | South Africans with Black and Indian Heritage vs. South Africans of European Heritage                     | 3.28            | 0.073           | 117             |
| US University Samples                                             | US Samples from Highly Selective Universities vs. a Primarily First-Generation College Student University | 0.53            | 0.467           | 653             |
| Testing Diff.<br>without<br>Translation                           | US vs. Hong Kong                                                                                          | 4.43            | 0.036           | 816             |
| Self-Report<br>Responsibilism                                     | US vs. Hong Kong                                                                                          | 19.25           | < 0.001         | 663             |
| "Culture-Report"<br>Responsibilism                                | US vs. Hong Kong                                                                                          | 15.37           | < 0.001         | 663             |

Note: This table checks the assumptions of independent samples *t* tests for whether variance is different between groups. When variance differs significantly between groups (as in the last three rows), some researchers recommend adjusting the degrees of freedom in the statistical test. That can change the significance values. After doing that for the tests in the last three rows, the differences between the US and Hong Kong remained significant (*Ps* < 0.001).

**Table S36B. Testing for equality of variance using Levene's test (part 2).**

| Items                                                                                                                                                                                                                                                                                                                         | <i>F</i> | <i>P</i> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|----------|
| I often have the feeling that my relationships with others are more important than my own accomplishments.                                                                                                                                                                                                                    | 0.002    | 0.966    |
| It is important that I do my job better than others.                                                                                                                                                                                                                                                                          | 0.31     | 0.578    |
| Winning is everything.                                                                                                                                                                                                                                                                                                        | 1.02     | 0.315    |
| We should keep our aging parents with us at home.                                                                                                                                                                                                                                                                             | 0.20     | 0.658    |
| The well-being of my coworkers is important to me.                                                                                                                                                                                                                                                                            | 19.55    | < 0.001  |
| To me, pleasure is spending time with others.                                                                                                                                                                                                                                                                                 | 5.02     | 0.026    |
| I feel good when I cooperate with others.                                                                                                                                                                                                                                                                                     | 0.12     | 0.729    |
| It often happens that the interests of my group coincide with my own interests.                                                                                                                                                                                                                                               | 0.29     | 0.589    |
| I often do "my own thing."                                                                                                                                                                                                                                                                                                    | 1.09     | 0.298    |
| I have subdued my personal opinions in order to be part of my group.                                                                                                                                                                                                                                                          | 12.67    | < 0.001  |
| I enjoy being unique and different from others in many ways.                                                                                                                                                                                                                                                                  | 2.28     | 0.133    |
| One should live one's life independently of others.                                                                                                                                                                                                                                                                           | 9.03     | 0.003    |
| You are working in a company for two years, and one day you meet a co-worker from a different department that you've never met before. You've been working on a sales account that could be a big accomplishment. The co-worker asks about the account.<br>I would feel comfortable telling this co-worker about the account. | 6.17     | 0.014    |
| If my parents have a nice little thing that happened to them today (for example, they saw a very pretty flower or their co-worker complimented them), my parents would tell me the little story.                                                                                                                              | 1.13     | 0.290    |
| Imagine you live in an apartment with four roommates that you didn't know before you moved in. One day, two friends of your roommate come over. Your roommate says, "My room isn't big enough for both of my friends. Can one sleep in your room?"<br>I would happily agree.                                                  | 1.05     | 0.307    |

Note: This table checks the assumption (of independent samples *t* tests) that variance is the same between groups. When variance differs significantly between groups, a common recommendation is to adjust the degrees of freedom in the *t* test. Five items showed significant evidence for unequal variance. [Table S13A](#) and [S13B](#) report results for those items after adjustment. Supplemental Section 38 discusses this process in more detail. Participants were 119 college students in the US and 76 in China. Sample sizes differ slightly across scales because we randomly assigned participants to scales.

## Supplementary References

54. Eriksson, K. *et al.* Perceptions of the appropriate response to norm violation in 57 societies. *Nat. Commun.* **12**, 1481 (2021).
55. Bettache, K., Chiu, C. & Beattie, P. The merciless mind in a dog-eat-dog society: neoliberalism and the indifference to social inequality. *Curr. Opin. Behav. Sci.* **34**, 217–222 (2020).
56. Talhelm, T. Hong Kong liberals are WEIRD: Analytic thought increases support for liberal policies. *Pers. Soc. Psychol. Bull.* **44**, 717–728 (2018).
57. Adler, N. E., Epel, E. S., Castellazzo, G. & Ickovics, J. R. Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychol.* **19**, 586–592 (2000).
58. Johnson, T., Kulesa, P., Cho, Y. I. & Shavitt, S. The relation between culture and response styles: Evidence from 19 countries. *J. Cross-Cult. Psychol.* **36**, 264–277 (2005).
59. Baumgartner, H. & Steenkamp, J.-B. E. Response styles in marketing research: A cross-national investigation. *J. Mark. Res.* **38**, 143–156 (2001).
60. Gardiner, G. & Funder, D. Misgivings about measurement invariance. *Eur. J. Personal.* **38**, 889–895 (2024).
61. House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W. & Gupta, V. *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies*. (Sage, 2004).
62. Daniels, M. A. & Greguras, G. J. Exploring the nature of power distance: Implications for micro- and macro-level theories, processes, and outcomes. *J. Manag.* **40**, 1202–1229 (2014).
63. Schwartz, S. H. & Sagiv, L. Identifying culture-specifics in the content and structure of values. *J. Cross-Cult. Psychol.* **26**, 92–116 (1995).
64. Li, L. M. W., Adams, G., Kurtiş, T. & Hamamura, T. Beware of friends: The cultural psychology of relational mobility and cautious intimacy. *Asian J. Soc. Psychol.* **18**, 124–133 (2015).
65. Cohn, A., Maréchal, M. A., Tannenbaum, D. & Zünd, C. L. Civic honesty around the globe. *Science* **365**, 70–73 (2019).
66. Gelman, A., Shor, B., Bafumi, J. & Park, D. Rich state, poor state, red state, blue state: What's the matter with Connecticut? *Q. J. Polit. Sci.* **2**, 345–367 (2008).
67. Heine, S. J., Buchtel, E. E. & Norenzayan, A. What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychol. Sci.* **19**, 309–313 (2008).
68. Eid, M., Lischetzke, T., Nussbeck, F. W. & Trierweiler, L. I. Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C (M-1) model. *Psychol. Methods* **8**, 38–60 (2003).
69. Hu, L. & Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* **6**, 1–55 (1999).
70. MacCallum, R. C., Browne, M. W. & Sugawara, H. M. Power analysis and determination of sample size for covariance structure modeling. *Psychol. Methods* **1**, 130–149 (1996).
71. Bentler, P. M. & Bonett, D. G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol. Bull.* **88**, 588–606 (1980).
72. Chen, F. F. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model.* **14**, 464–504 (2007).
73. Rutkowski, L. & Svetina, D. Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educ. Psychol. Meas.* **74**, 31–57 (2014).
74. Nye, C. D. & Drasgow, F. Effect size indices for analyses of measurement equivalence: understanding the practical importance of differences between groups. *J. Appl. Psychol.* **96**, 966–980 (2011).
75. Luong, R. & Flake, J. K. Measurement invariance testing using confirmatory factor analysis and alignment optimization: A tutorial for transparent analysis planning and reporting. *Psychol. Methods* **28**, 905–924 (2022).

76. Welzel, C., Brunkert, L., Kruse, S. & Inglehart, R. F. Non-invariance? An overstated problem with misconceived causes. *Sociol. Methods Res.* **52**, 1368–1400 (2021).
77. Snijders, T. A. & Bosker, R. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. (Sage Publications, 2011).
78. Maas, C. J. M. & Hox, J. J. Sufficient sample sizes for multilevel modeling. *Methodology* **1**, 86–92 (2005).
79. Jak, S., Oort, F. J. & Dolan, C. V. A test for cluster bias: Detecting violations of measurement invariance across clusters in multilevel data. *Struct. Equ. Model. Multidiscip. J.* **20**, 265–282 (2013).
80. Kim, E. S., Cao, C., Wang, Y. & Nguyen, D. T. Measurement invariance testing with many groups: A comparison of five approaches. *Struct. Equ. Model. Multidiscip. J.* **24**, 524–544 (2017).
81. Marsh, H. W. *et al.* What to do when scalar invariance fails: The extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychol. Methods* **23**, 524–545 (2018).
82. Marsh, H. W., Hau, K.-T. & Wen, Z. In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Struct. Equ. Model.* **11**, 320–341 (2004).
83. Cohen, L., Manion, L. & Morrison, K. Tests. in *Research methods in education* 563–585 (Routledge, London, UK, 2017).
84. Hinton, P. R., McMurray, I. & Brownlow, C. *SPSS Explained*. (Routledge, East Sussex, UK, 2014).
85. Oyserman, D. The lens of personhood: Viewing the self and others in a multicultural society. *J. Pers. Soc. Psychol.* **65**, 993–1009 (1993).
86. Funder, D. C. & Ozer, D. J. Evaluating effect size in psychological research: Sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* **2**, 156–168 (2019).
87. Van de Vliert, E. The global ecology of differentiation between us and them. *Nat. Hum. Behav.* **4**, 270–278 (2020).
88. Dunn, A. M., Heggstad, E. D., Shanock, L. R. & Theilgard, N. Intra-individual response variability as an indicator of insufficient effort responding: Comparison to other indicators and relationships with individual differences. *J. Bus. Psychol.* **33**, 105–121 (2018).
89. Roth, P. L., Bobko, P., Van Iddekinge, C. H. & Thatcher, J. B. Social media in employee-selection-related decisions: A research agenda for uncharted territory. *J. Manag.* **42**, 269–298 (2016).
90. Leiner, D. J. Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys. in *Survey Research Methods* vol. 13 229–248 (2019).
91. Rammstedt, B., Kemper, C. J. & Borg, I. Correcting Big Five personality measurements for acquiescence: An 18-country cross-cultural study. *Eur. J. Personal.* **27**, 71–81 (2013).
92. Fischer, R. Standardization to account for cross-cultural response bias: A classification of score adjustment procedures and review of research in JCCP. *J. Cross-Cult. Psychol.* **35**, 263–282 (2004).
93. Rammstedt, B. & Farmer, R. F. The impact of acquiescence on the evaluation of personality structure. *Psychol. Assess.* **25**, 1137 (2013).
94. Albarracin, D., Jones, C. R., Hepler, J. & Li, H. Liking for action and the vertical/horizontal dimension of culture in nineteen nations: Valuing equality over hierarchy promotes positivity towards action. *Interam. J. Psychol.* **51**, 335–343 (2018).
95. Booyen, F., Guvuriro, S. & Campher, C. Horizontal and vertical individualism and collectivism and preferences for altruism: A social discounting study. *Personal. Individ. Differ.* **178**, 110856 (2021).
96. Talhelm, T. Vertical collectivism does not measure what it claims to measure. *Curr. Res. Ecol. Soc. Psychol.* (R&R).
97. Roberts, B. W., Walton, K. E. & Viechtbauer, W. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychol. Bull.* **132**, 1–25 (2006).
98. Pick, C. M. *et al.* Family still matters: Human social motivation across 42 countries during a global pandemic. *Evol. Hum. Behav.* **43**, 527–535 (2022).
99. Roberts, B. W. & DelVecchio, W. F. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *Psychol. Bull.* **126**, 3–25 (2000).



100. Link, P. *An Anatomy of Chinese: Rhythm, Metaphor, Politics*. (Harvard University Press, 2012).
101. Harati, H. & Talhelm, T. Cultures in water-scarce environments are more long-term oriented. *Psychol. Sci.* **34**, 754–770 (2023).
102. Murray, D. R. & Schaller, M. Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture. *J. Cross-Cult. Psychol.* **41**, 99–108 (2010).
103. Adams, G. The cultural grounding of personal relationship: Enemyship in North American and West African worlds. *J. Pers. Soc. Psychol.* **88**, 948–968 (2005).
104. Cross, S. E., Bacon, P. L. & Morris, M. L. The relational-interdependent self-construal and relationships. *J. Pers. Soc. Psychol.* **78**, 791–808 (2000).
105. Schweinsberg, M. *et al.* The pipeline project: Pre-publication independent replications of a single laboratory's research pipeline. *J. Exp. Soc. Psychol.* **66**, 55–67 (2016).
106. Curran, P. J., West, S. G. & Finch, J. F. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol. Methods* **1**, 16–29 (1996).
107. Fischer, R. & Karl, J. A. A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Front. Psychol.* **10**, (2019).
108. Schumacker, R. E. & Lomax, R. G. *A Beginner's Guide to Structural Equation Modeling*. (Routledge, New York, NY, 2015).
109. Little, T. D. *Longitudinal Structural Equation Modeling*. (Guilford press, New York, NY, 2013).
110. Leung, K. & Bond, M. H. The impact of cultural collectivism on reward allocation. *J. Pers. Soc. Psychol.* **47**, 793 (1984).
111. Schmitt, D. P., Allik, J., McCrae, R. R. & Benet-Martínez, V. The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *J. Cross-Cult. Psychol.* **38**, 173–212 (2007).
112. Gosling, S. D., Rentfrow, P. J. & Swann, W. B. A very brief measure of the Big-Five personality domains. *J. Res. Personal.* **37**, 504–528 (2003).