AMERICAN PSYCHOLOGICAL ASSOCIATION

American Psychologist

ISSN: 0003-066X

2021, Vol. 76, No. 6, 997–1012
https://doi.org/10.1037/amp0000868

A Machine Learning Model of Cultural Change: Role of Prosociality, Political Attitudes, and Protestant Work Ethic

Abhishek Sheetal[1] and Krishna Savani[2]
[1] Division of Strategy, International Business, and Entrepreneurship, Nanyang Business School, Nanyang Technological University
[2] Division of Leadership, Management, and Organisation, Nanyang Business School, Nanyang Technological University

What attitudes, values, and beliefs serve as key markers of cultural change? To answer this question, we examined 221,485 respondents from the World Values Survey, a multiwave cross-country survey of people's attitudes, values, and beliefs. We trained a machine learning model to classify respondents into seven waves (i.e., periods). Once trained, the machine learning model identified a separate group of 24,611 respondents' wave with a balanced accuracy of 77%. We then queried the model to identify the attitudes, values, and beliefs that contributed the most to its classification decisions, and therefore, served as markers of cultural change. These included religiosity, social attitudes, political attitudes, independence, life satisfaction, Protestant work ethic, and prosociality. Although past research in cultural change has discussed decreasing religiosity and increasing liberalism and independence, it has not yet identified Protestant work ethic, political orientation, and prosociality as values relevant to cultural change. Thus, the current research points to new directions for future research on cultural change that might not be evident from either a deductive or an inductive approach. This research illustrates that the abductive approach of machine learning, which focuses on the most likely explanations for an outcome, can help generate novel insights.

Public Significance Statement
This research found that in recent years, people around the world have been becoming less religious and more liberal in their social attitudes and political orientation. People have been valuing independence more, although there appears to be a decline in the value of independence in the last few years. The extent to which people emphasize hard work, thrift, and prosociality has also declined in recent years.

Keywords: cultural change, gradient boosting, machine learning, World Values Survey

Supplemental materials: https://doi.org/10.1037/amp0000868.supp

Recent research in psychology has made substantial progress at documenting and predicting cultural change (Varnum & Grossmann, 2017). An emerging conclusion from this research is that cultural practices have been moving toward greater individualism over time. For example, the use of first-person pronouns has been increasing over time (Twenge et al., 2013), as has been the general use of words related to independence and individualism (Greenfield, 2013; Grossmann & Varnum, 2015; Twenge et al., 2012). These shifts can be traced to ecological changes, such as decreases in pathogen prevalence (Varnum & Grossmann, 2016) and to sociostructural shifts, such as those from rural to urban societies, agriculture to commerce, joint to nuclear households, and toward fewer children, more education, and more wealth (Greenfield, 2016; Grossmann & Varnum, 2015; Santos et al., 2017).

Much of the past research on cultural change has focused on sociostructural factors, probably because these variables are readily available over long periods and across many countries. Some research has utilized multicountry multiwave surveys to

Editor's note. This article is part of the special issue "Psychological Perspectives on Cultural Change" published in the September 2021 issue of *American Psychologist*. Michael E. W. Varnum and Igor Grossmann served as guest editors of the special issue, with Lillian Comas-Díaz as advisory editor.

Abhishek Sheetal https://orcid.org/0000-0002-4585-5358
Krishna Savani https://orcid.org/0000-0002-6934-1917

We thank Sylvia Chin, Ee-Hwee Lau, Kevin Nanakdewa, Elizabeth Demissie Degefe, and Xi Zou for helpful feedback on earlier drafts.

Correspondence concerning this article should be addressed to Krishna Savani, Division of Leadership, Management, and Organisation, Nanyang Business School, Nanyang Technological University, 50 Nanyang Avenue S3-01C-76, Singapore 639798, Singapore. Email: ksavani@ntu.edu.sg

Abhishek Sheetal

document the pattern of cultural change in people's attitudes and values. For example, researchers have used the World Values Survey (WVS; Haerpfer et al., 2020) to argue that people in most countries are moving toward greater individualism, that is, emphasizing self-expression, emphasizing friends over family members, and emphasizing the value of independence in children (Grossmann & Varnum, 2015; Hamamura, 2012; Inglehart & Baker, 2000; Santos et al., 2017). Many cultures are also moving toward a greater emphasis on modern, postindustrial, secular-rational values rather than traditional values; however, this shift is not as stark as the increase in self-expression (Inglehart & Baker, 2000).

Although it has been very productive, nearly all past research on cultural changes in attitudes, values, and beliefs suffers from a few limitations. First, out of the dozens of attitudes, values, and beliefs assessed in large multicountry multiwave surveys (e.g., the WVS), researchers typically focused on a select few questions (e.g., four questions examined by Hamamura, 2012; or three questions analyzed by Santos et al., 2017). Past research has primarily focused on changes in individualism (Greenfield, 2013; Grossmann & Varnum, 2015; Hamamura, 2012; Inglehart & Baker, 2000; Santos et al., 2017; Twenge et al., 2012). Although it is possible that the studied values are indeed the most important cultural dimensions that are changing over time, it is also possible that other attitudes, values, and beliefs might be similarly strong markers of cultural change.

Second, researchers have typically assumed that past patterns of cultural change would persist into the future, but cultural change can sometimes be nonmonotonic. For example, divorce rates in the United States peaked around 1980 and have since been declining (Hamamura, 2012). More generally, traditional statistical analyses require researchers to specify how they expect a given variable to change over time (e.g., linearly), but these assumptions might not hold in many cases given the complexity of cultural change (Erceg-Hurn & Mirosevich, 2008). Third, past research has rarely provided an independent test of the predictions of a statistical model of cultural change. Thus, it is possible that any given statistical model has overfit the available data and is unlikely to predict future data with similarly high accuracy.

In the current research, we seek to address these shortcomings by building a machine learning model of cultural change. First, unlike traditional analyses, machine learning models can handle a large number of variables; thereby, allowing us to use all attitudes, values, and beliefs that were measured in a multicountry, multiwave survey. Thus, we can objectively identify attitudes, values, and beliefs that serve as markers of cultural change from all the attitudes, values, and beliefs consistently measured in the survey. Second, unlike traditional regression-based analyses, machine learning models do not assume any functional form. Thus, they can model nonlinear changes in attitudes, values, and beliefs over time. Third, the accuracy of the machine learning model's predictions is tested in a subset of the data to which the model was never exposed (i.e., the *unseen data*); thereby, providing a metric of accuracy that is not susceptible to overfitting.

Method

All code is available at the online data repository for this project (https://osf.io/9yx8s; see file Code.zip). The procedure that we used is illustrated in Figure 1.

Dataset

We decided to use the WVS dataset (Haerpfer et al., 2020) because it is one of the most comprehensive international social science survey data sets currently available. It contains responses of 423,948 people from 104 different countries. The cross-sectional survey was conducted over seven periods (called *waves*): 1981–1984; 1989–1993; 1994–1998; 1999–2004; 2005–2009; 2010–2014; and 2017–2019. Thus, this dataset allows us to examine how people's attitudes, values, and beliefs have been changing over time in different parts of the world. Past research on cultural change has used this dataset (Hamamura, 2012; Inglehart & Baker, 2000; Santos et al., 2017). The WVS waves are highly irregular—the number of years covered by each wave and the gap between successive waves varies quite a bit. Thus, we treated each wave as a category rather than treating all waves as falling on a continuum.

The machine learning model that we used can automatically handle missing values. However, to minimize the proportion of missing values in the data, we only included 53 attitudes, values, and beliefs (*cultural variables* in short) that were asked in all seven waves in at least one country. A question assessing religious denomination (variable f025) included over 100 categorical response options, so we

Krishna Savani

recoded it into a binary variable (0 = no religious denomination, 1 = some religious denomination). Many of these 53 questions were not asked in several countries in a given wave. Once again, to minimize the proportion of missing data, we only included country-wave combinations in which all 53 cultural questions were administered.

This procedure resulted in a dataset with 246,096 respondents from 83 countries. Overall, 3.00% of the values were missing in this dataset. The list of variables and country-wave combinations included is available in the online data repository. The categorical variables e003, e004, and f034 were one-hot coded (i.e., a new variable was created for each response option), which resulted in a total of 61 variables. A list of all variables included is uploaded on the online data repository (see file WVS_variables_used.xlsx). A list of all country-subregion-wave combinations included in our dataset is also uploaded in the online repository (see file Country_wave_combinations.xlsx).

Only one country (Mexico) was included in all seven waves; one (Argentina) in six waves; three (Brazil, Chile, and the United States) in five waves; eight in four waves; 12 in three waves; 25 in two waves; and 33 in only one wave. Thus, for the vast majority of the countries, there was insufficient data to model how individuals' attitudes, values, and beliefs changed across different waves. Therefore, we combined data from multiple countries within each International Organization for Standardization [ISO] subregion (ISO, 2006). There was data from Sub-Saharan Africa and Latin America and the Caribbean for all seven waves; Eastern Asia, Southern Asia, and Southern Europe for six waves; Northern America, South-eastern Asia, and Eastern Europe for five waves; Western Asia, North Europe, and Australia and New Zealand for four waves; Northern Africa for two waves; and Central Asia for only one wave. Following past research (Inglehart & Baker, 2000; Schwartz, 2006), we assumed that countries in a given subregion are culturally similar. We did so because countries within a given region typically experience a similar ecology (Oishi & Graham, 2010) and because national boundaries are often a result of historical accidents (e.g., arbitrary colonial boundaries; Englebert et al., 2002). National boundaries are dichotomous, but cultural variation over geographical regions is more of a gradient (Cohen, 2001).

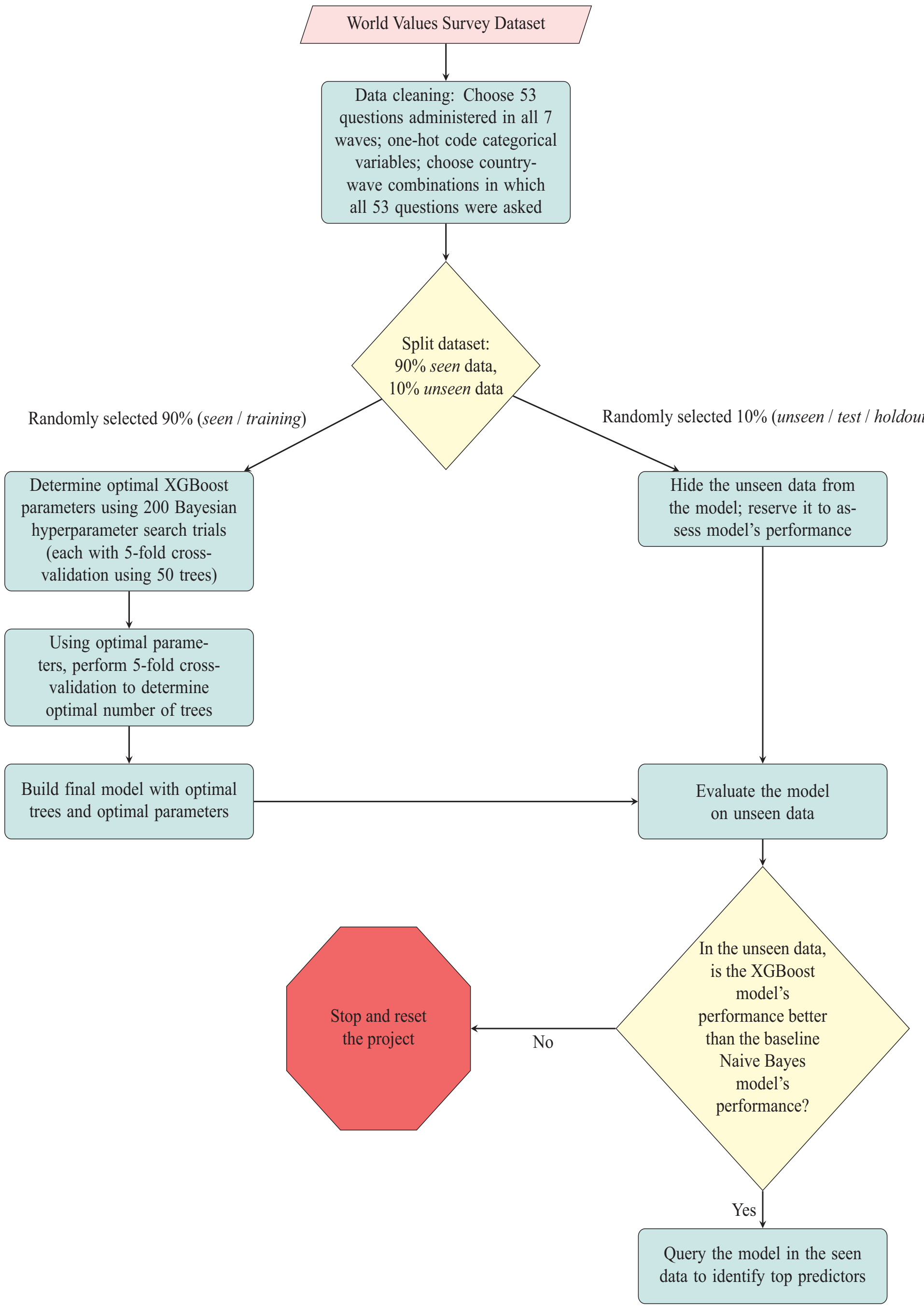
We used the holdout technique to test the model's predictions. That is, we randomly selected 24,611 respondents (10% of the data) and designated them as the *unseen sample*. The remaining 221,485 respondents were used to build the machine learning model (i.e., the *training data*). The machine learning model was not exposed to the unseen sample at any stage until it was finalized. Thus, the model's accuracy was assessed on these "new respondents" that the model had never seen previously. This process ensured that the training data and the unseen data were separated in all steps of the modeling process.

Building the Machine Learning Model

Our dataset was large enough that it would take weeks for it to be analyzed using the original CPU-based *random forest* machine learning method (Breiman, 2001) to complete the analysis on a 128-thread CPU. We performed the model training on the computer's graphics card (Nvidia GTX 1070) to reduce the model building time (Strigl et al., 2010). We used the *XGBoost* (Chen & Guestrin, 2016) package in *R*, which implements gradient boosted decision trees. The base random forest model is a special case of the more generic XGBoost model (XGBoost Developers, 2020a). XGBoost has been used in past research in psychology (e.g., Goretzko & Bühner, 2020) and economics (e.g., Kleinberg et al., 2018). Unlike random forest, XGBoost automatically handles missing values during the process of building a model; that is, it creates parallel paths for each possible missing value (e.g., assuming that a missing binary variable had a value of 0 in one path and a value of 1 in another path), and learns the optimal default path over successive iterations (Chen & Guestrin, 2016). In contrast, a standard random forest model would drop all observations with even a single missing value (Breiman, 2001).

We built three different models using the training data. In Model 1, we ignored the country and subregion that each participant belonged to, and instead focused on the wave in which each participant was sampled. The model learned to classify each participant in one of seven classes, representing the seven waves. In Model 2, we used a hierarchical structure with waves nested within subregions. This more

Figure 1
Illustration of the Machine Learning Procedure



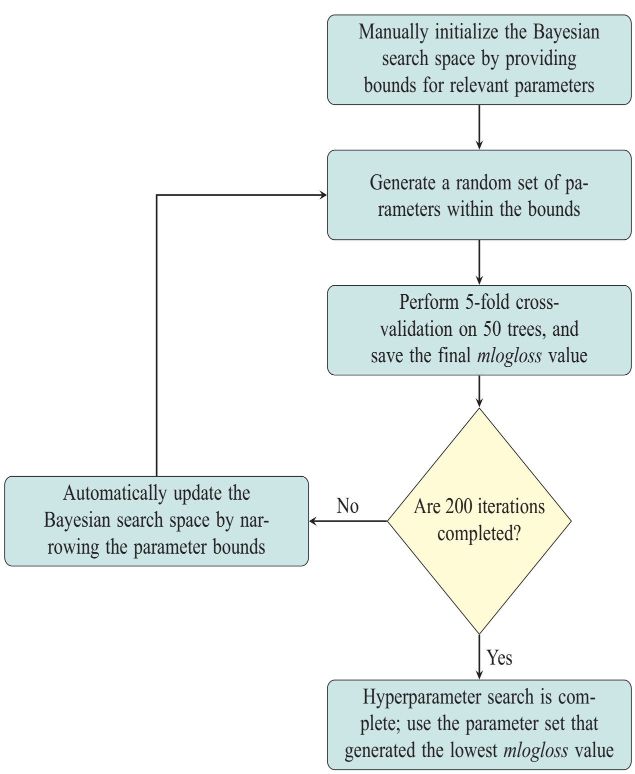
Note. See the online article for the color version of this figure.

complex model learned to classify each participant's wave along with their subregion (66 classes representing all wave-subregion pairs in our dataset). In Model 3, we treated waves as nested within countries. Thus, the model learned to classify each participant's wave and their country (179 classes representing all wave-country pairs in our dataset). As all models classified participants into one of many classes, we used the *multiclass-log-loss* loss function (also known as the *categorical cross-entropy* loss function; Zhang & Sabuncu, 2018).

Models trained to classify observations often accurately classify observations from the majority class but perform poorly when classifying observations from minority classes (Guo et al., 2008). As there was an uneven number of respondents in each wave (Model 1), each wave-subregion combination (Model 2), and each wave-country combination (Model 3), we assigned *class weights* to each observation. The *class weights* variable required by XGBoost must be a vector of N values, where N is the number of classes in the model (e.g., a vector of size 7 in Model 1). For each respondent in the training data, we divided the proportion of respondents belonging to the least represented class in the training data by the proportion of respondents belonging to the class to which this respondent belonged. Class weights were not used when the model was tested on the unseen sample. Note that class weights are distinct from the participant weights that the WVS provides (variable s017), which indicate whether each respondent's demographics characteristics are overrepresented or underrepresented in the survey relative to the population. It is not possible to simultaneously use class weights and participant weights in XGBoost.

XGBoost uses a three-step process to build the model. The first two steps identify the optimal model parameters, and the third step builds the actual machine learning model. In the first step, instead of using the default parameters in the XGBoost package, we conducted a hyperparameter search to identify a relatively superior set of XGBoost parameters (see Figure 2 for an illustration of the procedure). There are three methods for hyperparameter search. The first method is *random search*, in which the algorithm randomly draws a value of each parameter from the provided range across several iterations and picks the parameter set that generated the lowest loss value. The second method is *grid search*, in which the parameter values are stepped in increments over the parameter range until the entire space is exhausted. The third is *Bayesian search*, which uses information about how the loss value changed across all previous parameter combinations to guess the next set of parameters. We conducted a hyperparameter optimization search using the *BayesOpt* library (Martinez-Cantin, 2014), which can be applied to any parameter optimization problem. BayesOpt used fivefold cross-validation in each iteration; that is, 80% of the data was designated as the *training data*, which was used to identify a set of parameters, and the remaining 20% of the data was designated as the *validation data*, which was used to test the parameters. Across 200 iterations, the function experimented with various combinations of seven model parameters to identify the optimal set of parameters that generates the lowest multiclass log-loss in the validation data. Several other parameters were fixed. See Supplementary Materials Table 1 for details about the parameters.

Figure 2
Illustration of the Bayesian Hyperparameter Search Procedure



Note. See the online article for the color version of this figure.

In the second step, we took the optimal hyperparameter values determined in the first step and conducted a second parameter search procedure to determine the optimal number of trees that would minimize the loss function. Once again, we used fivefold cross-validation. The model started with one tree and assessed the loss value in the validation data. The model then kept increasing the number of trees across successive iterations up to a maximum of 2,000 trees. We set early stopping at five, which meant that if the validation data's loss value did not decrease appreciably after five iterations, the most recent number of trees was selected as the optimal parameter. In the third step, we took the parameters identified in the first two steps and built the final model. We repeated each of these three steps for each of our three models.

Results

Accuracy in Classifying Participants in the Unseen Sample

Once each model was finalized, we tested its performance in the unseen sample. Note that the model was never exposed to the unseen sample, and thus, the unseen sample serves as a set of "new participants" that the model has to classify based on each participants' attitudes, values, and beliefs. Table 1 reports the summary accuracy statistics.

The first five accuracy metrics reported in Table 1 are applicable to the model as a whole. The *no information accuracy* refers to the baseline accuracy rate if the model predicted that all participants in the unseen sample belonged

Table 1
Accuracy Statistics of the Three XGBoost Models in the Unseen Sample

Statistic	Model 1	Model 2	Model 3
No information accuracy	23.7%	5.7%	1.43%
Naïve Bayes accuracy	29.87% [29.30, 30.45]	13.30% [12.88, 13.73]	15.53% [15.08, 15.98]
XGBoost accuracy	59.18% [58.56, 59.80]	51.53% [50.90, 52.16]	50.49% [49.87, 51.12]
Kappa	50.55%	50.22%	50.18%
Area under the curve (AUC)	90.29%	96.54%	97.54%
Average sensitivity/recall	61.89%	51.56%	49.15%
Average specificity	92.96%	99.24%	99.72%
Average precision	58.37%	55.96%	48.68%
Average F1	59.83%	53.28%	48.79%
Average balanced accuracy	77.42%	75.40%	74.44%

Note. Numbers in parentheses refer to 95% confidence intervals.

to the most common class (e.g., Wave 7 in Model 1). In other words, if a rogue Model 1 predicted that every observation belonged to the majority class (i.e., Wave 7), then its accuracy would be 23.7%. The Naïve Bayes accuracy provides yet another benchmark; this model used Bayes' theorem to predict participants' class based on their attitudes, values, and beliefs, with the assumption that each of these variables independently contributes to the prediction (i.e., that there are no interactions; Rish, 2001). This accuracy is in the 18–28% range across the three models. The XGBoost model's accuracy ranged from 50–60%, which was substantially higher than the two benchmarks. Cohen's κ refers to the extent to which the model's accuracy is higher than the no information accuracy; our κ values of around 50% would be categorized as *moderate* according to Landis and Koch (1977). The XGBoost model's area under the curve (AUC) was over 90% for all three models. The AUC refers to the probability that when presented with two randomly selected participants from different classes (e.g., participant #1 from Wave 1 and participant #2 from Wave 2), the model predicts that participant #1's chances of being from Wave 1 are higher than participant #2's chances, and vice-versa for Wave 2. For multiclass classifiers, as in the present case, the AUC is a more relevant metric than raw accuracy (Huang & Ling, 2005).

The remaining accuracy metrics are specific to each class. The detailed class-by-class statistics for all three models are uploaded on the online data repository (see file Accuracy_statistics.xlsx). The average statistic across all classes is reported in Table 1. The models' *sensitivity* or *recall*, which refers to the proportion of participants from a given class who were accurately classified as belonging to that class, indicates that about half the participants belonging to a given class were accurately classified as such. The models' *specificity*, which refers to the proportion of participants not belonging to a given class who were accurately classified as not belonging to that class, is very high, exceeding 90%. The models' *precision*, which refers to the proportion of participants who actually belonged to the predicted class overall participants classified as belonging to that class, was above just about 50% across the three models. The models' *F1* statistic, which is the harmonic mean of precision and recall, was again about 50%. The models' *balanced accuracy*, which is the average of sensitivity and specificity, was between 74 and 77%, well above the raw accuracy rate of 50–59%. The balanced accuracy is particularly relevant for imbalanced data, in which some classes are overrepresented, and others are underrepresented, as in the present case.

We submit that the moderate raw accuracy rate of 50–59% is a function of the number of classes that the model was trained to distinguish (seven in Model 1, 66 in Model 2, and 179 in Model 3). With fewer classes, the model's accuracy should improve. To illustrate, we trained another XGBoost model to distinguish respondents from just Waves 6 and 7, which resulted in a binary classification problem. This model's accuracy was 81.56% (95% confidence interval, CI [80.71, 82.40]), significantly higher than Model 1's accuracy in the seven-class classification problem.

Querying the Model

XGBoost includes an in-built feature importance method (XGBoost Developers, 2020b) that can be used to rank order the predictor variables in terms of their contribution to the model's classification (often referred to as *gain* or *impurity* or *gini*; see Breiman et al., 1984). Alternatively, it is possible to use a permutation-based method, which would shuffle the values of each predictor variable one at a time and assess the extent to which the model's loss value changed as a consequence (Altmann et al., 2010). Variables whose perturbations lead to a bigger change in the model's loss value are relatively more important predictors. We decided to use a permutation-based method because "Permutation-based importance methods are a reliable technique that does not suffer from the bias existing in Gini impurity which might inflate the importance of continuous and high-cardinality categorical variables" (Gómez-Ramírez et al., 2020, p. 4). Specifically, we used the model-agnostic permutation-based DALEX package (Biecek, 2018). Note that XGBoost automatically models both linear and nonlinear relationships, along with any number of complex interactions. Thus, even if the mean value of a variable stayed largely constant across waves, the variable could show up as a top predictor in the variable importance analysis if it interacted with other variables.

Table 2 presents the top predictors from the first model that was trained to classify which of seven waves participants were sampled in. The rankings of all 61 predictor variables (after one-hot coding) in terms of their contribution to distinguishing respondents across the seven waves globally (Model 1), within each subregion (Model 2), and within each country (Model 3) are uploaded on the online data repository (see file Top_Predictors_DALEX.xlsx). We discuss the top 10 variables that contributed the most to distinguishing respondents from all seven waves, as well as across successive waves, from Model 1 (see file Top_predictors_model1.xlsx). For comparison, we also rank-ordered the predictors based on XGBoost's in-built *feature importance* algorithm and found substantial overlap with the DALEX-based feature importance rankings (see Top_predictors_XGBoost_Gini.xlsx on Open Science Framework [OSF]).

To illustrate the pattern of changes in the top predictors across the seven waves in Figures 3 to 7, we ran a hierarchical linear model (HLM) for each variable featured in Table 2 using the *xtreg* command in STATA. Specifically, in the original WVS dataset, we regressed each predictor on six dummy variables indicating Waves 2 to 7, respectively (Wave 1 was treated as the reference category). We treated participants as nested within countries, thereby accounting for country effects. Given the paucity of waves within a number of countries, we used fixed slopes. As we accounted for country effects, the mean levels of the variables depicted in the graphs are not interpretable—only the trends over time are interpretable. The HLM code and results are uploaded on the online data repository (HLM_code_figures.pdf). As different variables were measured on different response scales, for ease of interpretation, in the graphs, we scaled all variables to range from 0 to 1. We reverse-scored variables as needed for ease of interpretation. The HLM model was run on the original variables provided in the WVS dataset (before scaling and reverse-scoring).

Religiosity

The most important set of predictors that helped distinguish respondents across all seven waves was associated with religiosity (variables f028, f063, f025, and a040).

Consistent with the secularization hypothesis (Bruce, 2002, 2011), there is a trend toward lower religiosity over time (see Figure 3).

Social Attitudes

The second most important set of predictors that helped distinguish respondents from all seven waves referred to social attitudes. The justifiability of homosexuality (variable f118) and abortion (variable f120) helped distinguish respondents across all seven waves. Several related constructs helped distinguish people from different pairs of waves: the justifiability of divorce (variable f121) and euthanasia (variable f122), and the importance of tolerance (variable a035; see Figure 4). Of these, the acceptance of homosexuality showed the sharpest increase over successive years, consistent with past research (e.g., Loftus, 2001). The overall trend is toward greater liberalization of attitudes in these domains (Inglehart & Baker, 2000).

Independence

The next most important predictor that helped distinguish respondents from all seven waves was the importance of independence in children (variable a029; see Figure 5). The ninth most important predictor assessed the importance of obedience in children (variable a042). Additionally, a variable assessing whether greater respect for authority in the near future would be a bad thing (variable e018) distinguished respondents from pairs of waves. These findings are consistent with the existing literature documenting consistent increases in individualism over time (Santos et al., 2017).

Protestant Work Ethic/Confucian Work Ethic

The next most important predictor that helped distinguish respondents from all seven waves was the importance of hard work in children (variable a030). This variable consistently increased from Waves 1 to 4, and then declined slightly from Waves 4 to 7 (see Figure 6). On a related note, the importance of thrift (variable a038), determination (variable a039), and responsibility (variable a032) in children contributed to distinguishing respondents between various pairs of waves. These constructs appear closely related to the Protestant work ethic (Weber, 1930) and the Confucian work ethic (Lim, 2003), both of which emphasize hard work, personal responsibility, and thrift. A related construct, whether less importance on work in the future would be a bad thing (variable e015), helped distinguish respondents from Waves 1 versus 2 and 2 versus 3. This variable gradually decreased from Wave 1 to Wave 7, likely indicating an increased emphasis on leisure and diverged from the item emphasizing hard work in children.

Life Satisfaction

The next most important predictor that helped distinguish respondents from all seven waves was financial satisfaction

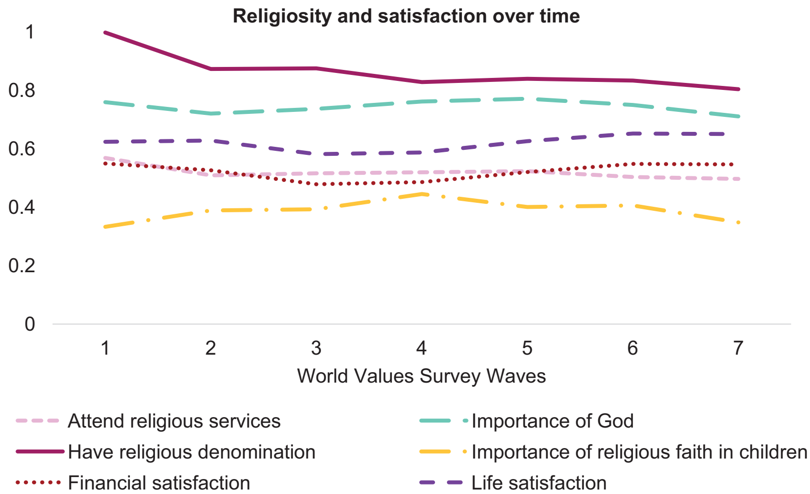
Table 2

Top 10 Attitudes, Values, and Beliefs That Helped Distinguish Respondents From All Seven Waves (Model 1), as Well as Those That Helped Distinguish Respondents From Each Successive Pairs of Waves (Also From Model 1)

Rank	All seven waves	Waves 1 vs. 2 (1981–1984 vs. 1989–1993)	Waves 2 vs. 3 (1989–1993 vs. 1994–1998)	Waves 3 vs. 4 (1994–1998 vs. 1999–2004)	Waves 4 vs. 5 (1999–2004 vs. 2005–2009)	Waves 5–6 (2005–2009 vs. 2010–2014)	Waves 6 vs. 7 (2010–2014 vs. 2017–2019)
1	Attend religious services (0.23)	Attend religious services (0.654)	Attend religious services (0.287)	Attend religious services (0.182)	Important in children: independence (0.168)	Justifiability of abortion (0.179)	Justifiability of abortion (0.306)
2	Justifiability of homosexuality (0.198)	Important in children: Hard work (0.408)	Importance of God (0.145)	Importance of God (0.135)	Attend religious services (0.156)	Justifiability of homosexuality (0.178)	Justifiability of homosexuality (0.245)
3	Justifiability of abortion (0.181)	Justifiability of homosexuality (0.36)	Greater respect for authority in future: Good or bad? (0.122)	Important in children: Obedience (0.101)	Important in children: Thrift (0.144)	Important in children: Independence (0.157)	Importance of God (0.185)
4	Importance of God (0.169)	Political orientation (0.322)	Political orientation (0.109)	Important in children: Independence (0.095)	Important in children: Religious faith (0.142)	Attend religious services (0.147)	Justifiability of accepting bribes (0.179)
5	Important in children: Independence (0.134)	Important in children: Independence (0.295)	Justifiability of homosexuality (0.101)	Political orientation (0.088)	Important in children: Obedience (0.141)	Importance of God (0.135)	Confidence in the armed forces (0.178)
6	Important in children: Hard work (0.133)	Religious denomination (0.287)	Less importance on work in future: Good or bad? (0.099)	Greater respect for authority in future: Good or bad? (0.082)	Important in children: Hard work (0.139)	Important in children: Thrift (0.13)	Justifiability of euthanasia (0.166)
7	Financial satisfaction (0.132)	Less importance on work in future: Good or bad? (0.249)	Justifiability of divorce (0.094)	Important in children: Religious faith (0.081)	Important in children: Unselfishness (0.138)	Important in children: Unselfishness (0.129)	Financial satisfaction (0.163)
8	Political orientation (0.13)	Confidence in the police (0.243)	Important in children: Obedience (0.086)	Important in children: Hard work (0.079)	Important in children: Determination and perseverance (0.137)	Important in children: Hard work (0.127)	Confidence in the parliament (0.162)
9	Important in children: Obedience (0.121)	Importance of God (0.242)	Life satisfaction (0.086)	Important in children: Tolerance (0.077)	Justifiability of homosexuality (0.125)	Important in children: Religious faith (0.12)	Justifiability of cheating on taxes (0.154)
10	Political action: Signing a petition (0.121)	Political action: Signing a petition (0.24)	Important in children: Hard work (0.084)	Important in children: Feeling of responsibility (0.077)	Importance of God (0.12)	Important in children: Determination and perseverance (0.117)	Attend religious services (0.147)

Note. Numbers in parentheses refer to the dropout loss associated with each predictor within that column—the bigger the number, the more important the variable.

Figure 3
Changes in Religiosity-Related Variables Over Time



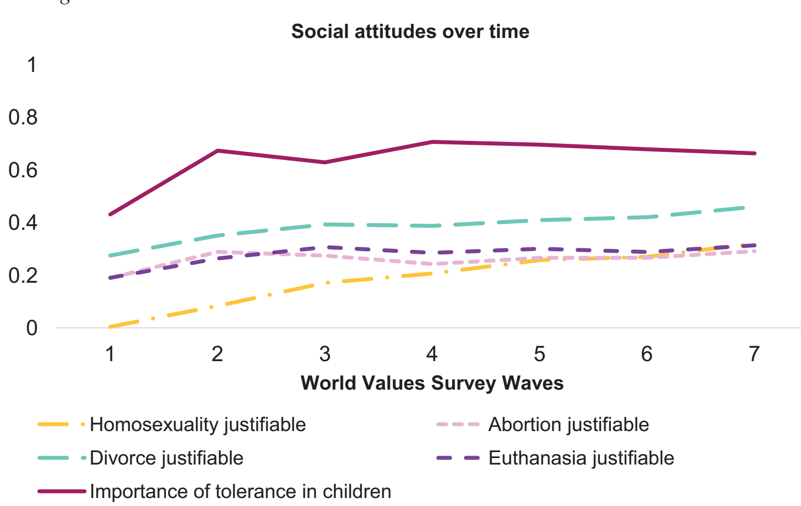
Note. Frequency of attending religious services was reverse-scored—higher numbers indicate greater frequency of attending religious services. See the online article for the color version of this figure.

(variable c006). A related variable, life satisfaction (variable a170), helped distinguish people between Wave 2 and 3. Both these variables declined from Waves 1 to 3, and then gradually increased (see Figure 3).

Political Attitudes

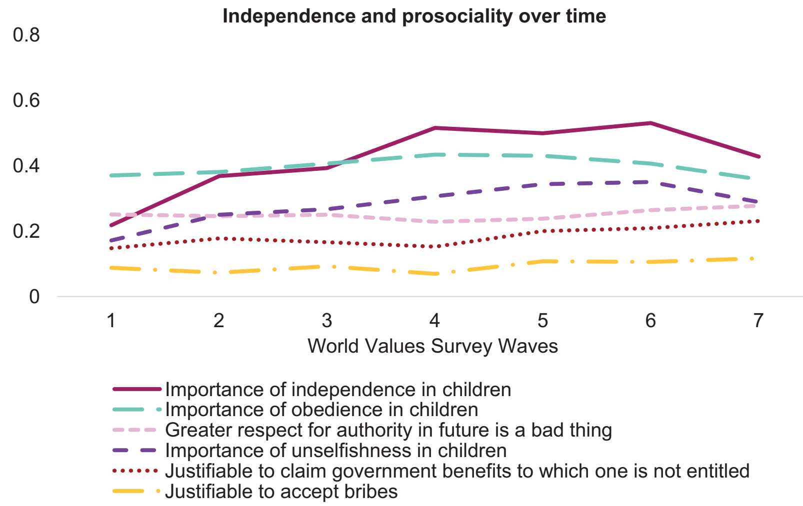
Next, political orientation (variable e033) helped distinguish respondents from all seven waves (see Figure 7). A related predictor referred to participants' tendency to engage in political action, that is, whether they have ever signed a petition, might do so in the future, or would never do so (variable e025). Respondents' political orientation became more liberal between Waves 1 and 2, and somewhat more liberal between Waves 6 and 7. Respondents' proclivity to sign a petition increased from Wave 1 and 2, then gradually declined until Wave 6, and then increased once again. Additional political attitudes helped distinguish respondents between various pairs of waves: confidence in the police (variable e069_06; Waves 1 vs. 2), and confidence in the armed forces and the parliament (variables e069_02 and e069_07; Waves 6 and 7).

Figure 4
Changes in Social Attitudes Over Time



Note. See the online article for the color version of this figure.

Figure 5
Changes in Independence and Prosociality Over Time



Note. See the online article for the color version of this figure.

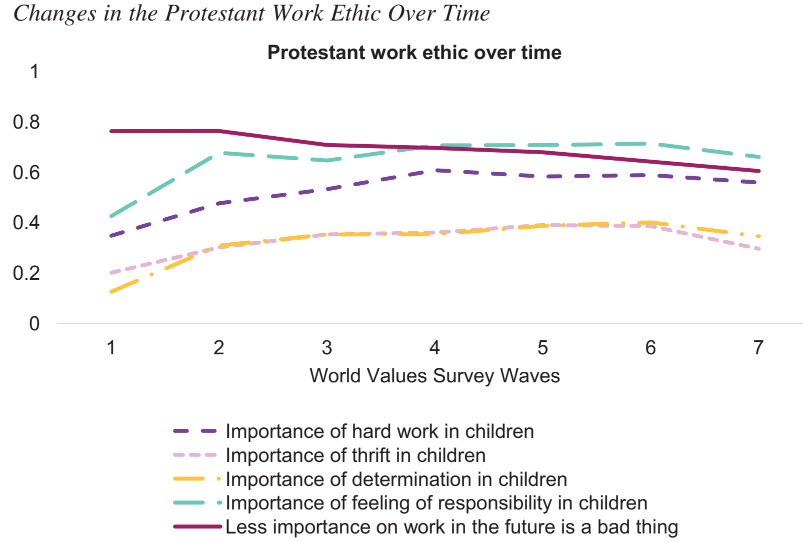
Prosociality

Three questions related to prosociality were relevant in distinguishing respondents between different pairs of waves, including the importance of unselfishness in children (variable a041), and the justifiability of accepting bribes (variable f117) and of cheating on taxes (variable f116). The importance of unselfishness increased from Wave 1 to Wave 6, but decreased from Wave 6 to Wave 7 (see Figure 5). The justifiability of claiming unauthorized government benefits and the justifiability of accepting bribes increased gradually from Waves 4 to 7. These findings indicate a curvilinear pattern, such that the importance of unselfishness reduced in recent years, whereas the justifiability of fraud increased.

Patterns of Cultural Change Within Different Sub-Regions

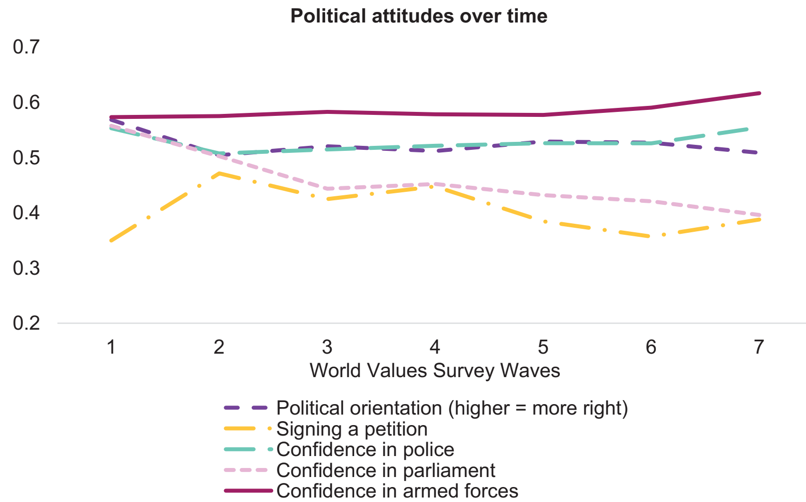
Next, we analyzed the variables that contributed the most to distinguishing respondents from different waves within each ISO subregion (from Model 2; see file Top_predictors_model2.xlsx on the OSF repository). Note that not all subregions were sampled in all waves. In Northern America, the key variables that distinguished respondents from different waves were religiosity, ethicality, attitudes toward homosexuality,

Figure 6
Changes in the Protestant Work Ethic Over Time



Note. See the online article for the color version of this figure.

Figure 7
Changes in Political Attitudes Over Time



Note. Signing a petition, confidence in police, and confidence in parliament have been reverse-scored so that higher numbers indicate greater willingness to sign a petition/more confidence in the police/the parliament/the armed forces. See the online article for the color version of this figure.

political attitudes, and Protestant work ethic. Life satisfaction and independence, which featured in the global Model 1, did not contribute to distinguishing North Americans from different waves. In Latin America and the Caribbean, the findings were fairly similar except that independence also contributed to distinguishing respondents from different waves.

In Northern Europe (i.e., the United Kingdom, the Nordic countries, and the Baltic states), the key features were Protestant work ethic, religiosity, social attitudes, independence, political attitudes, and a distinctive factor—national pride. An examination of the means across waves indicated that national pride increased from Wave 1 to Wave 5, but then decreased in Wave 6. The findings from Western Europe (i.e., French, German, and Dutch speaking countries) were fairly similar. National pride once again featured among the top 10 variables. It decreased from Wave 2 to Wave 3, but then gradually increased till Wave 7. In Southern Europe (i.e., Portugal, the Mediterranean countries, countries in the former Yugoslavia) and Eastern Europe (i.e., the remaining former Communist countries), religiosity, social attitudes, political attitudes, financial satisfaction, and independence contributed the most to distinguishing people across different waves.

In Northern Africa (i.e., countries on the Mediterranean coast, Sudan, and Western Sahara), religiosity, political attitudes, and Protestant work ethic contributed the most to distinguishing people across different waves. Health status featured as a key variable—it improved from Waves 4 to 6, but then declined in Wave 7. Independence, social attitudes, and life satisfaction did not contribute as much in Northern Africa, and prosociality did not feature in either Northern or Sub-Saharan Africa. In Sub-Saharan Africa, religiosity, political orientation, life satisfaction, and social attitudes were relevant. However, unlike the Americas and Europe, half of the top 10 predictors in Sub-Saharan Africa dealt with characteristics that are important for children. They covered the themes of Protestant work ethic and independence that featured in the global model. The importance of hard work, independence, responsibility, obedience, and thrift all increased from Wave 1 to Wave 5 or 6, and then declined in the last one or two waves. It appears that people's attitudes about child rearing were particularly relevant for distinguishing people across different waves in Sub-Saharan Africa.

In Western Asia (i.e., Turkey, and countries in the Caucasus and the Arabian peninsula), religiosity, social attitudes, financial satisfaction, political attitudes, Protestant work ethic, and respect for authority contributed the most to distinguishing people across different waves. In Southern Asia (i.e., Iran and the Indian subcontinent), religiosity, political attitudes, social attitudes, importance of obedience, and freedom of choice and control contributed the most to distinguishing respondents across different waves. Life satisfaction and prosociality did not feature. This was the only subregion in which the item assessing freedom of choice and control helped distinguish respondents across different waves. This variable decreased from Waves 2–3 to Wave 4, but then increased sharply till Wave 7. Thus, in recent decades, South Asians are experiencing more freedom of choice and control, which is associated with independence (Savani et al., 2008).

In South-eastern Asia (i.e., the region east of the Indian subcontinent and south of China), the key predictors were religiosity, Protestant work ethic, political attitudes, independence, and tolerance. Eastern Asia (i.e., China, Japan, Korea, Mongolia, and Taiwan) was one of the only two subregions in which religiosity was not the topmost factor that differentiated respondents across different waves. Instead, the most important factor was whether greater respect for authority would be a bad thing. This variable is on a downward trend over the years, indicating reducing respect for authority over time. A variable unique to Eastern Asia was willingness to fight for the country. This variable increased from Waves 1 to 2, but has been on a decline since. National pride also featured among the top 10 predictors. It increased between Waves 1 and 2, then decreased until Wave 5, and then increased again in Waves 6 and 7. The remaining predictors corresponded to religiosity, political attitudes, social attitudes, and importance of leisure. Australia and New Zealand represented the other subregion in which religiosity was not the top predictor. Instead, signing a petition, and justifiability of prostitution and homosexuality, contributed the most to distinguishing people across different waves. Other predictors included other political attitudes, religiosity, importance of authority, and emphasis on leisure. Life satisfaction and prosociality did not feature.

Patterns of Cultural Change Within Different Countries

In additional analyses, we examined the variables that contributed the most to distinguish different waves within each country (Model 3; see file Top_predictors_model3.xlsx on the OSF repository). Variables related to religiosity featured among the top 10 predictors for most countries; however, there was a large amount of between-country variability in terms of the variables most helpful in distinguishing respondents from different waves. Given the large number of countries, we do not discuss country-by-country results here. Further, given the paucity of waves in many countries, these results should be interpreted with caution. The rate of missing values varied substantially across the various country-wave combinations (see file MissingValues.xlsx on the OSF repository). Therefore, the accuracy statistics and variance importance results of Model 3 need to be interpreted with caution.

Lasso Regression

In further analysis, we ran a multinomial lasso regression (Tibshirani, 1996) in the training data to predict participants' wave based on all 61 variables, using 10-fold cross-validation to determine the appropriate lambda value. Lasso regressions automatically drop observations with even a single missing value, so this analysis was performed on 122,676 responses without any missing value across the training data and the unseen data. The lasso regression's accuracy in the unseen data was 44.21% (κ = 29.62%), lower than the XGBoost models' accuracy indicated in Table 1. The lasso regression identified the relevant predictors for distinguishing each wave from the baseline wave (i.e., Wave 1), but does not have an in-built feature to rank order the predictors in terms of their contribution to simultaneously distinguishing all seven waves. We again used the permutation-based DALEX package to rank order the predictors in terms of their contribution to distinguishing all seven waves in the lasso model.

The top 10 predictors were justification of homosexuality, political action: signing a petition, justification of abortion, importance of God, confidence in the parliament, justifiability of claiming undeserved government benefits, importance of obedience and unselfishness in children, and being an atheist (the complete list of variables is available in the Lasso_Predictors.xlsx file on OSF). Most of these variables were also identified as important predictors by the XGBoost model; however, the XGBoost model identified variables associated with the Protestant work ethic and life satisfaction, which were not identified by the lasso regression. The partial overlap between the two models is not surprising because the lasso regression makes a number of assumptions (e.g., sparsity; Hastie et al., 2015); whereas the XGBoost model does not make any assumptions (Chen & Guestrin, 2016).

Variables Not Contributing to Distinguishing Respondents From Different Waves

Of the 61 variables included in our analysis, only a minority contributed to distinguishing respondents from different waves, both globally (Model 1) and within each subregion (Model 2). A key question arises: Were the remaining variables largely constant across successive waves, or whether they did change across waves but were not as diagnostic of the waves in which respondents were sampled as the other variables? To answer this question, for each of the 61 variables, we computed the mean for each wave, and then computed the standard deviation of the mean across all waves. Next, we ranked the 61 variables in terms of their standard deviation. This information has been uploaded to the online data repository (see file WVS_variables_used.xlsx).

The top three variables with the highest variance across waves (i.e., importance of God, the justifiability of homosexuality, and financial satisfaction) all featured as top predictors in Model 1 and Model 2. The next two variables in the variance ranking—justifiability of euthanasia and abortion—did not contribute highly to distinguishing all seven waves (Model 1), but did contribute to distinguishing specific pairs of waves (Model 1), as well as waves within different subregions (Model 2). Several variables in the top third in terms of the variance ranking, including justifiability of prostitution and suicide, did not feature in either Model 1 or Model 2. Three variables—importance of independence in children, less importance on work in the future, and greater respect on authority in the future, scored relatively low in the variance ranking (rank 26, 31, and 37, respectively), but featured among the top 10 variables that distinguished people across waves in Model 1. Thus, the machine learning model was not simply picking up variables that varied the most across waves.

Discussion

In most research on cultural change in cultural psychology, researchers examine how the mean levels of a few theoretically driven variables changed across successive years (e.g., Greenfield, 2013; Grossmann & Varnum, 2015; Twenge et al., 2012). This is also the case with past research on cultural change using the WVS (e.g., Hamamura, 2012; Inglehart & Baker, 2000; Santos et al., 2017). In contrast, the machine learning method that we used considers all possible variables as relevant to cultural change, without being restricted by existing theories. This feature leaves open the possibility that the machine learning model can identify variables that are not currently theorized to be important to cultural change. Further, the machine learning model does not solely focus on the mean levels of variables across time—several variables that had low variance across waves were used by the machine learning model to identify respondents' wave, whereas a number of variables with high variance were not used. Of the cluster of constructs identified by the machine learning model, four (i.e., religiosity, social attitudes, independence, and satisfaction) have been well-studied in the literature on cultural change, whereas the remaining three (i.e., Protestant work ethic, political attitudes, and prosociality) appear to be relatively less examined in this literature.

The findings about changes in political attitudes are consistent with the general idea that the world is moving toward greater liberalism; however, past research has defined liberalism quite broadly to include a wide range of social attitudes and economic policies (Inglehart & Baker, 2000; Simmons et al., 2006). To our knowledge, cultural psychologists have not considered *political orientation* as a key feature of cultural change. We found that globally, people's political orientation is moving from center-right toward center-left, with a marked shift from 2010–2014 to 2017–2019. This finding seems counter to the fact that many Western democracies saw the ascendence of right-wing parties in the second decade of the 21st century. Perhaps part of the increase in liberalism might be a reaction to this ascendance. Visible examples of political activism in recent years, such as the 2020 Black Lives Movement and the 2021 U.S. Capitol riots, are consistent with the machine learning model's suggestion people's tendency to engage in political action has been increasing. The 2008 global financial crisis and the increased level of income inequality worldwide (Piketty, 2014) might have precipitated the trend toward more left-wing political ideologies and increased political action in the most recent decade. Given the economic havoc caused by the coronavirus disease 2019 (COVID-19) pandemic in 2020–2021, we would expect these trends to persist in the coming decade. Overall, these findings suggest that psychologists can fruitfully examine changes in political orientation and attitudes.

Although there is a substantial literature on the Protestant work ethic (Furnham, 1984; Weber, 1930) and the Confucian work ethic (Lim, 2003); although researchers have emphasized a key role of the Protestant work ethic during the reformation period and before (Andersen et al., 2017), to our knowledge, researchers studying cultural change have not considered the Protestant work ethic as a key construct varying over recent decades. The findings indicate that the importance of values related to the Protestant work ethic had been increasing worldwide until the early 21st century but have reached their peak and have begun declining in recent years. A potential explanation is that hard work has been a key means for economic mobility in most countries over most of the past four decades (e.g., Ali, 1992; Ferreira et al., 2012). However, economic mobility has declined in recent years (Katz & Krueger, 2017), which can explain the reduction in the importance of the Protestant work ethic in the most recent seventh wave of the WVS (2017–2019).

Prosociality is a major topic of research in psychology (e.g., Penner et al., 2005), sociology (e.g., Simpson & Willer, 2015), and economics (e.g., Bénabou & Tirole, 2006). Although extensive research has documented cultural similarities (e.g., Klein et al., 2015) and differences (Strombach et al., 2014) in generosity, to our knowledge, researchers have not systematically studied changes in prosociality over time. Our findings indicate a curvilinear pattern of change over time. The justifiability of accepting bribes and claiming unauthorized government benefits reached a floor around 1999–2004 and then started increasing. The importance of unselfishness in children gradually increased over time but declined in the most recent wave of the WVS. A similar curvilinear pattern is visible in the Gallup World Giving Index survey (Charities Aid Foundation, 2019), which asked respondents across a large number of countries to indicate whether they had donated to charity, volunteered their time, or helped a stranger in the past month. The mean level of the index across the world generally increased from 2010 to 2016 but has declined since. Perhaps increasing income inequality over recent years can explain this pattern (Côté et al., 2015). Future research can investigate this and other possible explanations for this curvilinear pattern.

Past research on cultural change has primarily focused on psychological variables that are changing similarly across the globe. The machine learning model identified many such variables, as discussed above, but it also identified culture-specific markers of cultural change. Religiosity was the dominant factor across the globe except for Eastern Asia, where reduced respect for authority was the most important marker of cultural change, and in Australia and New Zealand, where political action was the most important marker. National pride was relevant to cultural change in Northern Europe, Western Europe, and Eastern Asia; freedom of choice and control was particularly relevant in Southern Asia; and importance of hard work, independence, responsibility, and thrift in children in Sub-Saharan Africa. These findings suggest that researchers can fruitfully examine both culture-general and culture-specific aspects of cultural change.

In addition to identifying which psychological factors were key markers of cultural change, the present analysis identified variables that were not. In particular, generalized trust (variable a165), a key variable in sociology, economics, psychology, and political science, including in past research on cultural change (e.g., Hamamura, 2012), did not feature among the top 10 variables in either the global model or the subregion model. Neither did racism, that is, whether respondents would be comfortable with a person of a different race as a neighbor (variable a124_02), despite likely decreases in explicit racism over the years (Swim et al., 1995). And neither did xenophobia, that is, whether respondents would be comfortable with a foreigner as a neighbor (variable a124_06). Although people's health and possibly their happiness increased over the years with economic development and poverty reduction, people's self-related health status and happiness were not among the key markers of cultural change (except in Northern Africa). In contrast, their life satisfaction and financial satisfaction were important markers of cultural change.

Understandably, the WVS researchers designed the survey to measure a large number of different constructs to maximize the breadth of coverage, rather than measuring a narrower set of constructs using a large number of items that can be used to form a scale. Therefore, we had to analyze the data at the level of individual items rather than at the level of scale averages. Although most of our discussion focused on groups of thematically related items, it would be ideal if the current findings can be verified with external data sets that did measure the relevant constructs using multiitem scales over multiple decades in multiple countries. Till then, any generalization from individual items to theoretical constructs must be viewed as tentative.

The WVS sampled different countries in different waves. This inconsistency raises the concern that the changes in values across different periods are confounded with the sampling of different countries in different periods. This is a shortcoming of the global Model 1. However, it is less of a concern in Model 2, which assesses changes in values within each subregion, as countries within a subregion tend to be culturally similar (Inglehart & Baker, 2000; Schwartz, 2006). To address this shortcoming, when analyzing the results of Model 1, we plotted the change in values over successive waves while accounting country effects. The results depicted in the graphs can be interpreted as the expected change in values assuming that the same countries were sampled across all waves.

A key limitation of machine learning methods is that they inherit the limitations of existing large data sets (Zou & Schiebinger, 2018). There exist sufficiently large social science data sets, such as the WVS and the General Social Survey, on which machine learning analyses can be performed. However, such data sets exclusively contain individuals' responses to a large number of self-report questions. Although much can be gleaned from self-reports, self-reports have limited utility in cultural psychology (Heine et al., 2002). Much of the advancements in the field has come from the insight that culture is represented in patterns of cognition, emotion, and motivation that people might not have conscious access to (Adams & Markus, 2004; Kitayama, 2002), but which can be measured using behavioral tasks that assess implicit tendencies (Kitayama et al., 2009). However, large enough data sets of implicit tendencies do not exist, and compiling them would entail enormous costs. Thus, as of now, machine learning methods cannot be used to study implicit psychological tendencies.

The method that we used did not automatically group items into thematically related categories. However, psychologists are typically not interested in individual items but in the underlying constructs. Machine learning methods for dimensionality reduction exist, but in supervised learning (in which the model is trying to predict an outcome variable), the latent dimensions are typically constructed to accurately predict the outcome variable, not to represent coherent constructs (van der Maaten et al., 2009). Given that psychologists prioritize cogent explanations over predictive accuracy (Yarkoni & Westfall, 2017), less than ideal interpretability is a limitation of machine learning models. Unsupervised machine learning methods can be used for dimensionality reduction without prioritizing prediction, but once again, there is no requirement that the items feeding into the latent dimensions cohere together.

In conclusion, the machine learning approach to cultural change verified the significance of changes in independence, which has been a critical focus of this literature. The machine model's findings about the significance of religiosity and social attitudes are consistent with the sociological literature on cultural change. However, the machine learning approach also identified multiple other aspects of cultural change that have not received much attention, including Protestant work ethic, political orientation, political action, and prosociality. Thus, the current research indicates that theory-blind machine learning approaches can complement traditional theory-driven approaches to generate novel insights that researchers might miss otherwise (Bleidorn & Hopwood, 2019; Sheetal et al., 2020).

References

Adams, G., & Markus, H. R. (2004). Toward a conception of culture suitable for a social psychology of culture. In M. Schaller & C. S. Crandall (Eds.), *The psychological foundations of culture* (pp. 335–360). Erlbaum.

Ali, A. J. (1992). The Islamic work ethic in Arabia. *The Journal of Psychology*, *126*(5), 507–519. https://doi.org/10.1080/00223980.1992.10543384

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, *26*(10), 1340–1347. https://doi.org/10.1093/bioinformatics/btq134

Andersen, T. B., Bentzen, J., Dalgaard, C. J., & Sharp, P. (2017). Pre-reformation roots of the protestant ethic. *Economic Journal*, *127*(604), 1756–1793. https://doi.org/10.1111/ecoj.12367

Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *The American Economic Review*, *96*(5), 1652–1678. https://doi.org/10.1257/aer.96.5.1652

Biecek, P. (2018). Dalex: Explainers for complex predictive models in r. *Journal of Machine Learning Research*, *19*, 1–5.

Bleidorn, W., & Hopwood, C. J. (2019). Using machine learning to advance personality assessment and theory. *Personality and Social Psychology Review*, *23*(2), 190–203. https://doi.org/10.1177/1088868318772990

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC Press.

Bruce, S. (2002). *God is dead: Secularization in the West* (Vol. 3). Blackwell.

Bruce, S. (2011). *Secularization: In defense of an unfashionable theory*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199654123.001.0001

Charities Aid Foundation. (2019). *CAF World Giving Index 10th edition*. Retrieved February 15, 2021, from https://www.cafonline.org/about-us/publications/2019-publications/caf-world-giving-index-10th-edition

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Cohen, D. (2001). Cultural variation: Considerations and implications. *Psychological Bulletin*, *127*(4), 451–471. https://doi.org/10.1037/0033-2909.127.4.451

Côté, S., House, J., & Willer, R. (2015). High economic inequality leads higher-income individuals to be less generous. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(52), 15838–15843. https://doi.org/10.1073/pnas.1511536112

Englebert, P., Tarango, S., & Carter, M. (2002). Dismemberment and suffocation: A contribution to the debate on African boundaries. *Comparative Political Studies*, *35*(10), 1093–1118. https://doi.org/10.1177/001041402237944

Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, *63*(7), 591–601. https://doi.org/10.1037/0003-066X.63.7.591

Ferreira, F. H., Messina, J., Rigolini, J., López-Calva, L. F., Lugo, M. A., & Vakis, R. (2012). *Economic mobility and the rise of the Latin American middle class*. World Bank Publications. https://doi.org/10.1596/978-0-8213-9634-6

Furnham, A. (1984). The protestant work ethic: A review of the psychological literature. *European Journal of Social Psychology*, *14*(1), 87–104. https://doi.org/10.1002/ejsp.2420140108

Gómez-Ramírez, J., Ávila-Villanueva, M., & Fernández-Blázquez, M. Á. (2020). Selecting the most important self-assessed features for predicting conversion to mild cognitive impairment with random forest and permutation-based methods. *Scientific Reports*, *10*(1), 20630. https://doi.org/10.1038/s41598-020-77296-4

Goretzko, D., & Bühner, M. (2020). One model to rule them all? Using machine learning algorithms to determine the number of factors in exploratory factor analysis. *Psychological Methods*, *25*(6), 776–786. https://doi.org/10.1037/met0000262

Greenfield, P. M. (2013). The changing psychology of culture from 1800 through 2000. *Psychological Science*, *24*(9), 1722–1731. https://doi.org/10.1177/0956797613479387

Greenfield, P. M. (2016). Social change, cultural evolution, and human development. *Current Opinion in Psychology*, *8*, 84–92. https://doi.org/10.1016/j.copsyc.2015.10.012

Grossmann, I., & Varnum, M. E. (2015). Social structure, infectious diseases, disasters, secularism, and cultural change in America. *Psychological Science*, *26*(3), 311–324. https://doi.org/10.1177/0956797614563765

Guo, X., Yin, Y., Dong, C., Yang, G., & Zhou, G. (2008). On the class imbalance problem. *2008 Fourth International Conference on Natural Computation* (Vol. 4, pp. 192–201). IEEE. https://doi.org/10.1109/ICNC.2008.871

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., & Puranen, P. (2020). *World Values Survey: Round seven - Country-pooled datafile*. JD Systems Institute & WVSA Secretariat. https://doi.org/10.14281/18241.1

Hamamura, T. (2012). Are cultures becoming individualistic? A cross-temporal comparison of individualism-collectivism in the United States and Japan. *Personality and Social Psychology Review*, *16*(1), 3–24. https://doi.org/10.1177/1088868311411587

Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: The lasso and generalizations*. CRC Press. https://doi.org/10.1201/b18401

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, *82*(6), 903–918. https://doi.org/10.1037/0022-3514.82.6.903

Huang, J., & Ling, C. X. (2005). Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, *17*(3), 299–310. https://doi.org/10.1109/TKDE.2005.50

Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American Sociological Review*, *65*(1), 19–51. https://doi.org/10.2307/2657288

ISO (International Organization for Standardization). (2006). *ISO 3166-1: Codes for the representation of names of countries and their subdivisions–Pt. 1: Country codes*. Retrieved February 15, 2021, from https://www.iso.org/iso-3166-country-codes.html

Katz, L. F., & Krueger, A. B. (2017). Documenting decline in U.S. economic mobility. *Science*, *356*(6336), 382–383. https://doi.org/10.1126/science.aan3264

Kitayama, S. (2002). Culture and basic psychological processes—toward a system view of culture: Comment on Oyserman. *Psychological Bulletin*, *128*(1), 89–96. https://doi.org/10.1037/0033-2909.128.1.89

Kitayama, S., Park, H., Sevincer, A. T., Karasawa, M., & Uskul, A. K. (2009). A cultural task analysis of implicit independence: Comparing North America, Western Europe, and East Asia. *Journal of Personality and Social Psychology*, *97*(2), 236–255. https://doi.org/10.1037/a0015999

Klein, N., Grossman, I., Uskul, A. K., Kraus, A., & Epley, N. (2015). It pays to be nice, but not really nice: Asymmetric evaluations of prosociality across seven cultures. *Judgment and Decision Making*, *10*, 355–364.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, *133*(1), 237–293. https://doi.org/10.1093/qje/qjx032

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lim, V. K. (2003). Money matters: An empirical investigation of money, face and Confucian work ethic. *Personality and Individual Differences*, *35*(4), 953–970. https://doi.org/10.1016/S0191-8869(02)00311-2

Loftus, J. (2001). America's liberalization in attitudes toward homosexuality, 1973 to 1998. *American Sociological Review*, *66*(5), 762–782. https://doi.org/10.2307/3088957

Martinez-Cantin, R. (2014). BayesOpt: A Bayesian optimization library for non-linear optimization, experimental design and bandits. *Journal of Machine Learning Research*, *15*(1), 3735–3739.

Oishi, S., & Graham, J. (2010). Social ecology: Lost and found in psychological science. *Perspectives on Psychological Science*, *5*(4), 356–377. https://doi.org/10.1177/1745691610374588

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, *56*, 365–392. https://doi.org/10.1146/annurev.psych.56.091103.070141

Piketty, T. (2014). *Capital in the twenty-first century*. The Belknap Press of Harvard University Press. https://doi.org/10.4159/9780674369542

Rish, I. (2001). An empirical study of the naive Bayes classifier. *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, *3*(22), 41–46.

Santos, H. C., Varnum, M. E. W., & Grossmann, I. (2017). Global increases in individualism. *Psychological Science*, *28*(9), 1228–1239. https://doi.org/10.1177/0956797617700622

Savani, K., Markus, H. R., & Conner, A. L. (2008). Let your preference be your guide? Preferences and choices are more tightly linked for North Americans than for Indians. *Journal of Personality and Social Psychology*, *95*(4), 861–876. https://doi.org/10.1037/a0011618

Schwartz, S. (2006). A theory of cultural value orientations: Explication and applications. *Comparative Sociology*, *5*(2–3), 137–182. https://doi.org/10.1163/156913306778667357

Sheetal, A., Feng, Z., & Savani, K. (2020). Using machine learning to generate novel hypotheses: Increasing optimism about covid-19 makes people less willing to justify unethical behaviors. *Psychological Science*, *31*(10), 1222–1235. https://doi.org/10.1177/0956797620959594

Simmons, B. A., Dobbin, F., & Garrett, G. (2006). Introduction: The international diffusion of liberalism. *International Organization*, *60*(4), 781–810.

Simpson, B., & Willer, R. (2015). Beyond altruism: Sociological foundations of cooperation and prosocial behavior. *Annual Review of Sociology*, *41*(1), 43–63. https://doi.org/10.1146/annurev-soc-073014-112242

Strigl, D., Kofler, K., & Podlipnig, S. (2010). *Performance and scalability of GPU-based convolutional neural networks. 18th IEEE Euromicro Conference on Parallel, Distributed and Network-Based Processing* (pp. 317–324). IEE.

Strombach, T., Jin, J., Weber, B., Kenning, P., Shen, Q., Ma, Q., & Kalenscher, T. (2014). Charity begins at home: Cultural differences in social discounting and generosity. *Journal of Behavioral Decision Making*, *27*(3), 235–245. https://doi.org/10.1002/bdm.1802

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, *68*(2), 199–214. https://doi.org/10.1037/0022-3514.68.2.199

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B. Methodological*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012). Increases in individualistic words and phrases in American books, 1960-2008. *PLoS ONE*, *7*(7), e40181. https://doi.org/10.1371/journal.pone.0040181

Twenge, J. M., Campbell, W. K., & Gentile, B. (2013). Changes in pronoun use in American books and the rise of individualism, 1960¬2008. *Journal of Cross-Cultural Psychology*, *44*(3), 406–415. https://doi.org/10.1177/0022022112455100

Van Der Maaten, L., Postma, E., & Van den Herik, J. (2009). Dimensionality reduction: A comparative review. *Journal of Machine Learning Research*, *10*, 66–71.

Varnum, M. E. W., & Grossmann, I. (2017). Cultural change: The how and the why. *Perspectives on Psychological Science*, *12*(6), 956–972. https://doi.org/10.1177/1745691617699971

Varnum, M. E., & Grossmann, I. (2016). Pathogen prevalence is associated with cultural changes in gender equality. *Nature Human Behaviour*, *1*, 1–4.

Weber, M. (1930). *The Protestant ethic and the spirit of capitalism*. Routledge.

XGBoost Developers. (2020a). *Random Forests in XGBoost*. Retrieved February 15, 2021, from https://xgboost.readthedocs.io/en/latest/tutorials/rf.html

XGBoost Developers. (2020b). *Understand your dataset with XGBoost*. Retrieved February 15, 2021, from https://xgboost.readthedocs.io/en/latest/R-package/discoverYourData.html

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zhang, Z., & Sabuncu, M. R. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. *arXiv*. https://arxiv.org/abs/1805.07836

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—It's time to make it fair. *Nature*, *559*(7714), 324–326. https://doi.org/10.1038/d41586-018-05707-8

Received July 31, 2020
Revision received April 15, 2021
Accepted April 16, 2021 ■