

Global Social Tolerance Index and multi-method country rankings sensitivity

By: Stelios H Zanakis, William Newburry, [Vasyl Taras](#)

This is a post-peer-review, pre-copyedit version of an article published in *Journal of International Business Studies*. The definitive publisher-authenticated version:

Zanakis, S., Newburry, W., & Taras, V. (2016). Global social tolerance index and multi-method country rankings sensitivity. *Journal of International Business Studies*. 47(5): 480-495. doi: 10.1057/jibs.2016.5

is available online at: <https://doi.org/10.1057/jibs.2016.5>

Abstract:

Social tolerance refers to the extent of recognition and acceptance of differences, willingness to grant equal rights, and refraining from openly intolerant attitudes. Utilizing World Value Survey (WVS) data (56 countries, 83,000 usable respondents), we develop a Global Social Tolerance Index (GSTI) that incorporates gender, minority, immigrant, and religion tolerance dimensions. We develop this index using a multi-step data-driven procedure involving five data standardizations, seven weighting approaches, and five aggregation methods. They generate 124 replications for each country's index and rank, yielding a median overall position and several measures of rank sensitivity/robustness to different weighting, standardization, and aggregation approaches. It builds on prior social tolerance indexes in terms of scope and dimensionality, and avoids problems associated with equal or subjective weighting. The GSTI index/ rankings provide a tool for IB scholars to examine nations' overall tolerance or tolerance regarding each dimension relative to other external criteria. Our procedure can be used to develop other indexes and rankings of nations or organizations within a country or region. In practice, any such non-participatory method should always serve as a starting point to facilitate deliberations of experts and/or decision-makers for policy recommendations.

Keywords: indices and rankings | measurement and scale development issues | social tolerance | gender equality | immigration and minorities | religion

Article:

Introduction

Tolerance generally refers to the ability to endure, but its specific meaning differs by discipline. We employ tolerance in sociocultural terms, as "sympathy or indulgence for beliefs or practices differing from or conflicting with one's own" (Merriam-Webster Dictionary, 2010). Social tolerance is the extent of recognition and acceptance of differences, willingness to grant equal rights, and refraining from openly intolerant attitudes (Hodges, Green, & Gurevich, 2001). Given the cognitive nature of this definition, we seek to identify a social tolerance index not from country-level outcomes (e.g., minority percentages), but from citizens' perceptual views as measured in the WVS (WVS, 2005). Scholars have identified differences between artifacts and

associated perceptions stemming from individuals' perceptual biases (e.g., Elsbach, 2005). While measuring artifacts is common due to information availability, by tapping into underlying perceptions, we hope to achieve a more accurate representation of tolerance within a society and improve prior research by providing a tool for IB scholars to assist in empirically testing research questions regarding social tolerance.

The present study advances earlier work that compared countries in terms of social tolerance (e.g., Das, DiRienzo, & Tiemann, 2008) in several important ways. First, recognizing that tolerance is a multidimensional construct, we develop a multi-factor social tolerance index that allows for a more fine grained analysis of the various aspects of social tolerance. Second, we use a more elaborate re-sampling median-based computation methodology that yields more reliable estimates of social tolerance indexes and rankings. Third, we present alternative measures of ranking sensitivity or robustness. Fourth, we validate our overall index and each of our dimensional indexes against a set of criteria and explore the role of social tolerance in a larger nomological network – demonstrating that our composite index correlates reasonably higher with our external validation measures than the Das et al. (2008) index. For calculating all indexes in this study we consider only non-participatory unsupervised methods that require no inputs from expert opinions or from specialized external software, helping to reduce researcher bias in the index development process.

While a general empirical link has been demonstrated between social tolerance and certain economic growth measures (Das et al., 2008; Treisman, 2000), a comprehensive social tolerance index may assist academics, regulators, government officials, and analysts in better identifying global economic determinants and consequences of this important construct. For example, a better understanding of tolerance may help government officials to develop better policies to address issues such as the integration of immigrants into the local economy. Differentiation between countries with respect to gender-based tolerance may help explain the level of integration of women into a country's workforce, and by extension, influence a company's competitiveness in the global economy. Moreover, a multidimensional tolerance measure will better equip scholars to develop and test comparative cross-national theory regarding which tolerance dimensions impact particular variables of interest across different institutional settings. Having a reliable social tolerance measure will help IB scholars explore important questions such as how the tolerance of a country directly or indirectly impacts foreign direct investment location choices, expatriate assignment decisions, and strategic marketing and branding decisions in a host country.

Although developing and validating a multidimensional Global Social Tolerance Index (GSTI) is itself worthwhile, we also seek to improve rankings in other areas that have been criticized for methodological flaws or questionable simplifications (e.g., Bardhan & Klasen, 1999; Dijkstra, 2002; Permanyer, 2011). The most common ills are using weights assumed equal or elicited dubiously from "experts", and employing a single aggregation method (e.g., weighted averages) without examining rank sensitivity. The GSTI developed herein uses a data-driven multi-method approach which avoids these tribulations and is supplemented with rank robustness/ sensitivity measures. The method proposed here can be used in developing other rankings/indices for nations or organizations within a country.

Social Tolerance Research, Dimensions, and Indexes

Despite its importance to international business, social tolerance research has been limited in terms of its theoretical foundation for index components and its methodology. For example, Das et al. (2008) offered a national-level tolerance index using average responses to four 1999–2004 WVS questions regarding minorities and homosexuals. While providing a foundation for our work, their study is limited in the number of items used, the consideration of only one overall tolerance dimension, and its ranking development method, in particular weighting all items equally. Similarly, using the General Social Survey of US households, Hodges Persell et al. (2001) derived an index regarding tolerance towards African Americans and homosexuals.

While prior research provides tools for comparing social tolerance cross-nationally, these indexes and rankings suffer from common disadvantages, namely: limited scope, limited dimensionality, and equal weighting. We address these shortcomings by offering a new social tolerance index derived using a re-sampling-type methodology. The proposed index contains gender, minorities, immigration, and religion factors, four of the most common issues associated with tolerance globally. While other tolerance areas exist (e.g., homosexuality, social class, political, differently abled, age), we contend that these four factors largely represent the major tolerance issues impacting global business today. Additionally, the tolerance dimensions vary in their degree of integration of a focal group into society (e.g., from willingness to accept to ability to become full citizens), capturing differences in their current states of acceptance. Furthermore, we employ a more robust methodology to derive national rankings.

We use the 2005 WVS database from 56 countries with over 83,000 usable individual responses (ranging from 954 for New Zealand to 3051 for Egypt, averaging 1456 per country) to measure the four tolerance components. The survey was translated from English to the local language for administration, with appropriate adjustments made to ensure functional equivalence (see www.worldvaluessurvey.org for additional information). Note that one of the initial 57 countries from this WVS (Hong Kong) was omitted due to a large number of missing values and outliers for this country.

Thirteen items were chosen based on past literature to represent the components of GSTI. To facilitate interpretations, items were reverse coded, when appropriate, so that higher values imply more tolerance (as shown in all 13 items of the four dimensions below). Note that the WVS items have different response scales, which if used without adjustment could be problematic. This is evident in the religion items, where the first two are binary and the other two require a 1-to-5 response. For the dichotomous variables, aggregating individual-level data essentially creates a country-level item that indicates the percentage of people in a country that agree with a statement. After the responses were put on a common metric with higher scores indicating greater tolerance and averaged at the country level, we obtained comparable continuous scales presented below. As such, improved reliabilities were obtained using national averages vs the initial WVS respondents' data for the 13 questions chosen.

Gender

Granting women equal rights to men is a longstanding global issue (Adler, 2002; Bullough, Kroeck, Newburry, Kundu, & Lowe, 2012; Parboteeah, Hoegl, & Cullen, 2008). For example, women were not allowed to vote until 1920 in the US, 1944 in France, 1947 in Mexico, 1971 in Switzerland, 2006 in the UAE, and December 2015 in Saudi Arabia (although many other gender-based restrictions remain). Several organizations have compiled national gender inequality rankings, such as the United Nations Development Program's Genderrelated

Development Index and the World Bank's Gender Equality Index. Although criticized for their construction method and assumptions made in overcoming data gaps (cf., Bardhan & Klasen, 1999; Dijkstra, 2002), these indexes remain popular in academic research. Whereas the above indexes utilize secondary data, the GLOBE project surveyed respondents regarding perceived gender role equality (House, Hanges, Javidan, Dorfman, & Gupta, 2004). While making an important contribution, the GLOBE survey items reflect general gender role attitudes as opposed to tolerance. Moreover, the items assessed current practices ("as is") and prescription ("should be") for a society, vs individual attitudes (e.g., "my personal beliefs are").

The above gender equality measures examined observable behaviors and actual women participation rates (UN) or respondents' assessments of gender role practices (GLOBE). We examine tolerance of gender equality, an important theoretical complement to past studies. It is not just a women's societal position that matters, but also whether women are accepted in these positions (Newbury, Belkin, & Ansari, 2008). Our measures examine respondents' personal perceptions, focusing on tolerance towards women enjoying equal rights, rather than actual women's rights. As such, measuring this tolerance dimension helps address calls for greater IB research on gender (e.g., WAIB, 2015) by addressing an important component of gender bias in international business.

Respondents evaluated the following items (4-point scale: 1=agree strongly to 4=disagree strongly):

- (1) Men make better political leaders.
- (2) University is more important for a boy than for a girl.
- (3) Men make better business executives than women do.

Minorities

Globalization has brought differences in minority group treatment across societies into focus, as scholars examine marginalization's impact within the corporation (Price & Feinman, 1995). Similar to gender, equal minority group treatment may indicate an ability to operate according to expected global business norms. Stone, Hosoda, Lukaszewski, and Phillips (2008) examined some problems of current minority discrimination research, including questionable measures, non-representative samples, and experimental situations that provide "hints and cues that govern the participants' perceptions of his or her role and of the experimenter's hypotheses" (Rosenthal & Rosnow, 1991: 115), thus limiting research validity. Wrench (2011) further noted cross-country data comparability problems. Weldon's (2006) minority tolerance index was based on five equally weighted questionnaire items from the 1997 Eurobarometer survey of 15 European Union members; none of these items is similar to the WVS questions used herein.

Developing a cross-national minority rights index appears problematic given difficulties in consistently classifying what constitutes a minority cross-nationally. Our measure allows for a more comparable examination of minority tolerance by examining perceptions regarding multiple minority groups. We help pave the way for cross-national study of minority issues by establishing a construct for use both in theoretical development and empirical analysis. This may be particularly useful to IB scholars interested in the impact of discrimination against minorities on important HR topics such as the perceptions of marginalized individuals in host countries and their relationships with foreign firms (e.g., Newbury, Gardberg, & Sanchez, 2014).

Three items measured our minority tolerance component, the first two of which resemble the ones used by Das et al. (2008) from WVS data prior to 2005.

On this list are various groups of people. Could you please mention any that you would not like to have as neighbors? (Binary response reversed to: 0 = mentioned, 1= not mentioned):

- (1) *People of a different race.*
- (2) *Immigrants/foreign workers.*
- (3) *People who speak a different language.*

Immigration

Recent estimates suggest that the global total of migrant workers would equal the population of the fifth largest country, with the number expected to increase in coming decades (ILO, 2013). As such, immigrant worker tolerance is increasingly important to global business. Given a frequent concern regarding immigrant-related employment impacts, immigrant-related research often focuses on workforce issues. For example, DelCampo, Jacobson, Van Buren, and Blancero (2011) examined perceptions of immigrant and non-immigrant US-born Hispanics. Thus IB scholars would benefit from an index addressing immigrant tolerance as it would help in examining issues of how firms deal with immigrant workers in both their home and host countries. Given the prominence of NGOs in recent IB research, an immigrant tolerance index might also aid in studying the work of NGOs in addressing such issues.

However, little research has examined immigration-related tolerance. An exception is Weldon's (2006) study of ethnic minority tolerance that focused a section on how citizenship regimes can affect social and political tolerances using 1997 Eurobarometer survey items. We fill this gap by deriving an immigration factor within our overall GSTI. Note that this dimension contrasts with the immigration item within the minority dimension since the minority item focuses on immigrant worker tolerance, while this dimension examines achievement of citizenship to capture a nation's stance on immigration. The items within our index are:

In your opinion, how important should the following be as requirements for somebody seeking citizenship of your country? (3-point scale: 1= very important, 2= rather important, 3= not important):

- (1) *Having ancestors from my country.*
- (2) *Being born on my country's soil.*
- (3) *Adopting the customs of my country.*

Additional WVS items were considered but not supported by subsequent reliability analysis.

Religion

While the first three components of our tolerance index concern tolerance towards specific groups, the final component of our index addresses a factor often associated with a lack of tolerance towards multiple groups: religion. As such, it considers tolerance at a more broad-based institutional level. "All major religions, specifically Christianity, Judaism, Islam, Buddhism, and Hinduism, include teachings and rituals that emphasize certain work values and

attitudes” (Yeganeh, 2015: 585). Thus religion plays a strong role in forming norms of behavior in a society. In particular, strong religious beliefs have commonly been associated with conservative attitudes (Cukur, de Guzman, & Carlo, 2014) and the pursuit of “absolute moral standards” (Yeganeh, 2015: 587), which tend towards strict definitions of acceptable behavior. These beliefs have commonly been associated with various forms of tolerance such as attitudes restricting the role of women in society (Seguino, 2011) and the acceptance and rights of homosexuals (Jäckle & Wenzelburger, 2015). Religion has also been linked to attitudes towards immigrants (e.g., Foner & Alba, 2008), and religious particularism has been linked to racial prejudice in Europe (Ekici & Yucel, 2015). Strict beliefs such as these may ultimately extend to a lack of tolerance of outsiders in general.

It is recognized that many religions subscribe to a version of the “Golden Rule”, “to do unto others as you would wish them to do unto you”, which would suggest greater tolerance. Moreover, while resisted by many religious leaders, recent efforts by Pope Francis and others to take a more inclusive stance regarding diversity issues such as acceptance of homosexuals are noteworthy. However, given the above strong evidence in past research linking religion with inflexible beliefs and various forms of intolerance, a dimension incorporating religion seems appropriate to be included within our overall social tolerance framework. As religious institutions play influential roles in many emerging markets, including a religion dimension in our tolerance index may be useful to IB scholars examining how institutional environments impact business practices affecting foreign entrants into these markets. To incorporate this, we used the following four items from the same WVS:

- (1) *Generally speaking, do you think that the churches* in your country are giving adequate answers to: The problems of family life? (Binary response reversed to: 0= yes, 1= no)*
- (2) *Generally speaking, do you think that the churches* in your country are giving adequate answers to: The social problems facing our society? (Binary response reversed to: 0=yes, 1=no)*
- (3) *It would be better for [respondent’s country] if more people with strong religious beliefs held public office. (5-point scale: 1= strongly agree to 5= strongly disagree)*
- (4) *Politicians who do not believe in god are unfit for public office. (5-point scale: 1 =strongly agree to 5= strongly disagree)*

*In non-Christian societies substitute “religious authorities” for “churches”.

The first two items constituted Müller’s (2009) “Religious Influence Index”, while the latter two have been previously used to construct his “Religious Leadership Index”. Three other promising WVS questions (“Religious leaders should not influence how people vote in elections”, “Religious leaders should not influence government decisions”, and “Having as neighbors people of a different religion”) were considered for inclusion, but later dropped after data analysis revealed they worsened the internal fit of this factor.

Psychometric Properties of the Scale

To assess the basic psychometric properties of the scales, we first subjected our data to exploratory factor analysis. Based on the exploratory factor analysis, the factor structure was nearly perfect at both the individual and the national levels of analysis. Likewise, confirmatory

factor analysis results provided strong empirical support for the four-factor structure, with all the item loadings and goodness-of-fit statistics significantly exceeding the commonly accepted cut-off points. Further, we tested the internal reliability of our subscales. All of these met commonly accepted standards at both the individual and national levels of analysis. Specifically, the Cronbach's alphas, at the individual and national levels of analysis respectively, were 0.78 and 0.93 for the Gender, 0.72 and 0.91 for Minorities, 0.73 and 0.85 for Immigration, and 0.71 and 0.95 for Religion tolerance components.

Furthermore, to justify the aggregation of the individual WVS data and their use at the country level of analysis, we computed intra-class correlations (ICCs). As noted by Hanges and Dickson (2004), ICCs reported in organizational literature tend to be rather low, averaging only 0.12 and typically ranging from 0.05 to 0.20. Similarly low figures are provided in studies that specifically focus on cross-national comparisons (e.g., Au & Cheung, 2004; Dolan, Diez-Pinol, Fernandez-Alles, MartinPrius, & Martinez-Fierro, 2004; Lenartowicz & Roth, 2001). The figures for aggregate data on attitudes and cultural values are even less impressive. A reanalysis of Hofstede's data, probably the most known national-level dataset derived by aggregation of individual responses reported ICC(1) of 0.02–0.07 (Gerhart & Fang, 2005), which is close to the average of 0.04 reported by Hofstede (1980). A meta-analysis of the research that used Hofstede's approach to measuring culture reported ICC(1) of 0.10–0.30, depending on the value dimension (Steel & Taras, 2010), which is similar to the GLOBE figures reported by Hanges and Dickson (2004): ICC(1) between 0.22 and 0.27 depending on the scale. Our own calculations for the WVS items used to compute GSTI were comparatively better. For the composite scales, ICC(1) ranged from 0.42 for the Gender subscale to 0.94 for the Minorities subscale. The figures were lower for individual items ranging from 0.11 to 0.95 (0.61 average), but still noticeably better than the ICC and within/between country variance statistics reported in earlier similar studies, justifying aggregation of individual data to the country level.

GSTI Methodology and Development

WVS country averages were used to compute GSTI in all subsequent work. In order to justify the countrylevel aggregation, ANOVAs were computed on the 13 study items to generate F-statistics comparing the between country vs within country differences using the individual-level respondents data. For the 13 items, the between-country variance was very significant ($p < 0.001$), ranging from 7.1% and 34.5%, justifying the aggregation.

We also considered the issue of cross-national response style differences (Smith, 2004). Correcting for possible cross-cultural differences did not seem appropriate in our case for multiple reasons. First, since responses to all items in our study are closely related, centering (within-subject or within-group) will remove important variance and can undermine the data validity – distorting the data and hiding real differences pertaining to our constructs. Second, given the limited response scale ranges for our study items (in fact, five of the 13 items were binary questions), we do not expect to see extreme response bias. More importantly, for the yes/no items, true country-level data properties do not become manifested until the data are aggregated to generate country-level percentages. Finally, our multi-ranking procedures per country require a single national-level value per attribute; thus they cannot be applied to individual-level multi-response data. However, to eliminate possible threats to validity of our findings, we tested our data for acquiescence bias and extreme response styles (Marin, Gamba, & Marin, 1992) and compared rankings with and without corrections. The results revealed no

significant response style differences across the cultural groups represented in our sample, and the relative rankings of the countries did not change appreciably when response style corrections were used. Given the above reasons against using such corrections in our case, the results based on the data uncorrected for response style are reported here.

Not surprisingly, prior research into international differences in social tolerance, as well as other country rankings generally refrained from data transformations that could correct for response style. For example, Das et al. (2008) did not correct their social tolerance data for cross-cultural differences in response style. Within-subject standardization was applied by Hofstede to his IBM data (Hofstede, 1980) and a similar standardization was used in the GLOBE project (House et al., 2004), but the standardizations were made across all items in the survey which referred to a wide range of values and issues, not to the same issue as is the case in our study.

Using national averages for our analyses addresses concerns about the impact of differences in country sample sizes. The original dataset (56 countries, 13 items) contained 10.9% missing values, which were imputed using the widely used EM loglikelihood maximization algorithm (Dempster, Laird, & Rubin, 1977: 42,177 Google Scholar citations as of 5 August 2015). The algorithm may work reasonably well when the underlying normality assumption does not hold (as in our study) or data are not missing completely at random (MCAR) provided the missing mechanism does not depend on the observed complete data (Little, 1992; Pigott, 2001). In our case, Little's MCAR test, comparing the actual patterns of missing data with those expected if they were distributed as MCAR is plausible ($p \sim 0.10$, not significant), empirically justified by the majority of missing data coming from countries where a question was not administered rather than being the choice of individual respondents. No country's missing value estimates are outside their pre-imputation range, and none of the imputed value ranks are out of sync with their non-missing data ranks. A comparison of absolute mean differences of the 13 items before and after imputation shows that they are all small between 0.0% and 1.6%, the latter for the item: "churches give answers to problems of family life". This religion item had seven missing values compared to the immigration items where pre-post imputation differences are only 0.22–0.67% in spite of their larger number of fully imputed scores (see below). Additionally, the pre-post differences in variability of each item were also small. In general, imputation of missing values reduced the skewness of most items, as did the elimination of both index and rank outliers. The above explorations render additional confidence that missing values were reasonably estimated in this dataset. Thus we elected to keep imputed values in the tables with a clear indication regarding their imputation rather than omitting them all, which would have eliminated many major countries and would lead to an undesirably smaller number of responses (see discussion in conclusion section).

Note that immigration data were not collected in eleven countries (Britain, Colombia, France, Guatemala, Iran, Iraq, Japan, the Netherlands, New Zealand, Peru, and Russia), resulting in the scores for the Immigration Index to be fully imputed for these countries. Fewer full imputations occurred for the Religion Index (Britain, China, France, the Netherlands, and Russia), for the Minorities Index (Egypt, Iraq, and Japan), and none for the Gender Index. These scores are marked by "a" in Table 1 (see footnote) and caution should be used in interpreting these values.

Reliability analyses on national averages with or without missing values (i.e., prior to or after the EM imputations) validated all four factors. Cronbach's alphas at the national level are in the 0.90s, except for immigrants (0.85), and substantially higher than those with the WWS raw data (0.70s). Even when deleting the two countries with the largest number of missing values

(Colombia and Iraq), all four factors' reliability results remain equivalent. In light of these analyses, the 13 WVS items, although not exhaustive, appear reasonably valid and sufficiently robust as input for the GSTI.

We use a data-driven multi-stage approach to derive the GSTI. First, all data are converted to a common scale by performing five standardizations: percent of max, 0–1 range, Euclidean vector, percentile rank, and z-scores. Second, the data are weighted using seven weighting methods: equal weights, two modifications of Pawlak's (1991) rough sets AI, entropy with two transformations, and two hybrid square distance variations from non(ideal) TOPSIS targets (Hwang & Yoon, 1981). Lastly, five aggregation methods are used to obtain the composite index and rank for each country by combining each attribute weight and value according to: simple additive weighting, weighted product ratio, ordering by similarity to ideal solution TOPSIS (Hwang & Yoon, 1981), and weight free scale-invariant combination of the first two aggregations using two distribution origin location estimators (Muralidhar & Zanakis, 1992). This re-sampling-type process generates not just one index value and rank but up to 175 (=5×7×5) estimates per country – depending on certain method requirements for data standardization. Any choice of specific procedures cannot be claimed to be exhaustive or the best. For additional information on this methodology, see Zanakis and Rickling (2010) and Simpson (2011).

Our analyses produced 124 valid combinations of the above three procedural steps, each yielding an index and rank per country. Outliers were identified non-parametrically, due to lack of normality, and eliminated at $\alpha = 0.05$ (typically about 3% at a subindex level and 1% at the overall index level). For example at the overall median rank level, the 71 country outliers averaged a data outlier reduction of 1.2% with the largest number of outliers deleted (out of 124 replications) occurring for France (17), Italy (22), and India (20). Four countries had 1–5 outliers only, while the remaining 49 countries had no rank outliers. Similarly, for the overall median index, 132 outliers (1.9%) were deleted, primarily from Andorra (23), Jordan (17), Bulgaria (14), Norway (14), and Sweden (13); this notably includes the top three ranked countries because, as often observed in rankings, the top and bottom few countries form their own clusters of larger consecutive steps in an ordered list. Deleting outliers at the national data level tends to eliminate the combined effect of external “noises” within individual-level data as well as procedure-level “noises”.

Then, using the corresponding attributes, we obtain for each country the median rather than the mean of the generated cleaned sub-indexes as the most representative tolerance score for each dimension (reasons for using the median are explained later): Gender Equality, Minorities, Immigrants, and Religion (Table 1, Columns 1–4), along with their correlations (see table bottom). All subindex correlations are, as expected, positive and highly significant ($p < 0.01$). A more detailed discussion of these results is provided later.

Our methodology differs from others cited in three important ways: (1) we utilize more items to quantify each of four tolerance dimensions, even compared to researchers using WVS data (c.f., Das et al., 2008; Knack & Keefer, 1997; Treisman, 2000); (2) we obtain data-driven weighted factor scores rather than simple averages of items loading in each factor (cf., Das et al., 2008); and (3) instead of the customary single index, we employ combinations of procedures to generate large re-samples of indexes and ranks per country, resulting in confidence-based assessments of index and rank, along with sensitivity analysis.

The aggregate GSTI is obtained using the four tolerance sub-indexes to develop a composite. For each country, the median of the 124 method combinations of index values

described earlier is used as that country's GSTI (Column 8). The commonly used benchmark index of equal weighted sub-indexes (Column 7) and the corresponding benchmark ranks are also shown (Columns 5–6). Examining the lack of proximity between the median (sub)index and the equally weighted benchmark reveals the extent of their overall (dis)agreement: The overall GSTI matched the benchmark ranks in 19 of the 56 countries. The minority tolerance index is the most similar to the equally weighted benchmark (mean absolute percent difference, MAPD =0.3%), while the immigrant and religion indexes are the most different (MAPD =4.4% and 4.3%). As an external validation of our GSTI, we added (Column 9) the GT H&M Index of Das et al. (2008) for 31 countries common to both studies. The correlation between these two indexes (0.62) is significant, even though only two of their four questions (both on minorities) matched our 13 items, and our construct is more theoretically and empirically justifiable.

As expected, the seven weighting methods produce rather dissimilar weights from each other and from the benchmark's equal weights of 0.25, resulting in the following overall weights for the four sub-indexes: Religion (0.282), Immigrants (0.268), Gender (0.236), and Minorities (0.214). The smallest weight for minorities is influenced by its items' smallest data variability and high negative skewness (about -1.18), while the religion items' approximate data symmetry and largest variability impact its higher weight. Notably, method weights are influenced by input data variability at the index level (standard deviation) and at the attribute item level (coefficient of variation due to markedly different means and shapes). For each of the 124 combinations, our procedure generates an index for each country and corresponding rank among the 56 countries. We report for each country the median of the ranks generated rather than the (single) rank of the median index calculated. In other words, for each country, ranks have their own distribution independently of the median index and then we separately calculated the median of these ranks. We chose the median index and the median of ranks generated for several more reasons: Lack of normality for most countries, and sensitivity of a mean to outliers and to non-constant variance across the scores;¹ furthermore, a country may have index outliers and no rank outliers (like Andorra) or vice versa. The median response has been used in Delphi assessments to arrive at a consensus of expert opinions in research and practice widely since the mid1950s (see, e.g., Linstone & Turoff, 2002).

In Table 2, we use the difference of median minus benchmark rank to measure the extent of disagreement with the equally weighted benchmark rank, found to be + three positions at most. A negative difference implies that our approach improves that country's ranking compared with that of equal weighting, while a positive difference worsens it. Out of 56 countries, the median index improves the ranking of 19, worsens 18, and leaves 19 the same. To quantify each country's rank sensitivity or robustness, we used four rank dispersion measures:

- (a) Range of the 95% non-parametric confidence interval for a country's median rank.
- (b) Range of all method combination ranks for the country.
- (c) Same range of index ranks standardized relative to the country's median rank.
- (d) Percent of ranks within the highest frequency five-rank group steps.

Outliers had been removed prior to the analysis. One can also visualize the extent of these rank sensitivities or robustness as the range of these steps around the main diagonal (the perfect median rank location nicely fitted in this case).

Note that a higher value in the first three measures implies more sensitivity, but the opposite holds for the last measure. The undisputed robust countries are those with the best

possible value for each measure, namely Sweden and Norway, followed by Andorra, and Switzerland. Andorra is the most consistently ranked country out of the 56, placed 3rd by all 124 procedures (with no rank outliers). All 13 of its item data ranks are single digits except one (which ranked 14). Notably, the most sensitive country rankings are concentrated around the middle, including Japan at 24th and South Korea at 26th (with the two biggest rank ranges and five-rank majorities <50%), Guatemala at 40th and Trinidad and Tobago at 22nd. As expected, the most robust countries gravitate towards the top and bottom ranks (with few exceptions).

Table 1. Global Social Tolerance Indexes and their correlations

Country	Gender Median Index (1)	Minorities Median Index (2)	Immigrant Median Index (3)	Religion Median Index (4)	GSTI Equal Weight Benchmark Rank (5)	GSTI Median Rank (6)	GSTI Equal Weight Benchmark Index (7)	GSTI Median Index (8)	Das et al. GT_H&M Index ^b (9)
Andorra	0.828	0.990	0.726	0.913	3	3	0.903	0.899	
Argentina	0.678	0.982	0.596	0.585	12	12	0.735	0.727	0.724
Australia	0.742	0.925	0.577	0.694	10	9	0.761	0.757	
Brazil	0.631	0.924	0.573	0.295	23	25	0.621	0.583	
Britain	0.704	0.890	0.525 ^a	0.531 ^a	17	16	0.683	0.666	0.752
Bulgaria	0.618	0.762	0.491	0.668	21	20	0.652	0.650	0.289
Burkina Faso	0.263	0.860	0.229	0.223	41	43	0.390	0.333	
Canada	0.810	0.969	0.505	0.610	11	11	0.749	0.731	0.838
Chile	0.535	0.903	0.392	0.476	26	27	0.589	0.558	0.658
China	0.452	0.765	0.409	0.420 ^a	35	33	0.518	0.497	0.324
Colombia	0.627	0.917	0.363 ^a	0.254	33	34	0.549	0.497	
Cyprus	0.655	0.765	0.456	0.599	22	22	0.635	0.618	
Egypt	0.000	0.187 ^a	0.119	0.097	55	55	0.071	0.045	0.501
Ethiopia	0.776	0.794	0.230	0.177	36	38	0.499	0.408	
Finland	0.759	0.823	0.623	0.631	13	10	0.733	0.733	0.762
France	0.921	0.599	0.695 ^a	0.793 ^a	6	6	0.781	0.775	0.833
Georgia	0.332	0.694	0.171	0.040	46	48	0.297	0.191	
Germany	0.817	0.862	0.644	0.719	5	5	0.790	0.789	0.778
Ghana	0.295	0.647	0.139	0.045	50	51	0.267	0.176	
Guatemala	0.641	0.957	0.291 ^a	0.073	37	40	0.495	0.349	
India	0.262	0.408	0.149	0.514	45	45	0.324	0.311	0.394
Indonesia	0.431	0.506	0.209	0.047	48	48	0.285	0.207	0.232
Iran	0.187	0.409	0.188 ^a	0.188	52	49	0.225	0.208	0.711
Iraq	0.013	0.000 ^a	0.000 ^a	0.049	56	56	0.000	0.000	
Italy	0.705	0.848	0.621	0.535	15	15	0.699	0.696	0.637
Japan	0.495	0.490 ^a	0.539 ^a	0.817	25	24	0.598	0.590	
Jordan	0.181	0.157	0.229	0.128	54	52	0.149	0.154	0.000
Malaysia	0.266	0.503	0.111	0.111	51	52	0.230	0.191	
Mali	0.029	0.682	0.012	0.021	53	54	0.163	0.061	
Mexico	0.580	0.872	0.213	0.457	34	35	0.539	0.482	
Moldova	0.471	0.721	0.526	0.446	32	29	0.550	0.540	0.139
Morocco	0.375	0.646	0.122	0.010	49	51	0.274	0.122	0.508
Netherlands	0.822	0.882	0.597 ^a	0.645 ^a	9	8	0.764	0.760	0.974
New Zealand	0.771	0.946	0.621 ^a	0.668	7	6	0.780	0.776	
Norway	1.000	0.971	0.825	1.000	2	2	0.995	0.993	
Peru	0.645	0.938	0.374 ^a	0.329	27	30	0.583	0.540	0.515
Poland	0.546	0.844	0.406	0.460	29	28	0.575	0.548	0.462
Romania	0.546	0.780	0.673	0.245	30	30	0.572	0.529	0.231
Russia	0.351	0.710	0.293 ^a	0.202 ^a	42	42	0.384	0.350	0.429
Rwanda	0.402	0.446	0.219	0.470	43	41	0.379	0.372	
Serbia	0.601	0.741	0.577	0.442	24	24	0.604	0.592	0.437
Slovenia	0.694	0.812	0.552	0.718	14	14	0.717	0.715	
South Africa	0.542	0.830	0.165	0.202	39	40	0.434	0.352	0.347
South Korea	0.472	0.519	0.636	0.641	28	26	0.578	0.579	0.512
Spain	0.868	0.937	0.377	0.776	8	12	0.767	0.728	0.795
Sweden	0.922	1.000	1.000	0.890	1	1	1.000	1.000	1.000
Switzerland	0.865	0.930	0.784	0.658	4	4	0.843	0.841	
Taiwan	0.535	0.808	0.708	0.547	19	18	0.668	0.663	
Thailand	0.451	0.559	0.106	0.109	47	46	0.294	0.224	
Trinidad and Tobago	0.757	0.954	0.519	0.335	20	22	0.660	0.621	
Turkey	0.491	0.604	0.384	0.337	38	37	0.455	0.440	0.003
Ukraine	0.406	0.796	0.557	0.406	31	32	0.550	0.534	0.344
United States	0.712	0.871	0.532	0.518	18	18	0.678	0.666	0.752
Uruguay	0.693	0.946	0.383	0.641	16	19	0.686	0.661	
Vietnam	0.488	0.474	0.275	0.497	40	37	0.433	0.428	0.409
Zambia	0.478	0.609	0.241	0.204	44	43	0.378	0.342	
Correlations (2)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(3)	0.774	0.582							
(4)	0.742	0.460	0.803						
(5)	-0.911	-0.731	-0.891	-0.881					
(6)	-0.895	-0.694	-0.907	-0.899	0.995				
(7)	0.936	0.789	0.903	0.871	-0.977	-0.973			
(8)	0.906	0.723	0.922	0.908	-0.979	-0.983	0.991		
(9)	0.638	0.508	0.459	0.616	-0.703	-0.691	0.633	0.618	

Note: *Italics* indicates the minimum value of an index and **Bold** indicates the maximum value.
^aWARNING: Scores based on imputed data.

^bThe Das et al. Index values shown in Table 1 reflect a reversed stretched scale such that 1 = maximum tolerance and 0 = minimum tolerance.

GSTI Validation Against Theoretically Relevant External Criteria

The validation assessed how the composite GSTI and its sub-indexes fit into a larger nomological network of theoretically related constructs. We clarify that these constructs are not competing comprehensive social tolerance measures, but rather variables that should theoretically relate to tolerance. In addition to validating GSTI against the Das et al. (2008) findings, we also use data from Hofstede (1980, 2001), GLOBE (House et al., 2004), Schwartz (1999, 2008), Leung and Bond (2004), and archival databases of Freedom House, the World Bank, and the United Nations. When possible, mid-2000s indices were used to achieve temporal precedence with the WVS data. The GLOBE study also reported separate scores for their major cultural subgroups for a few countries (e.g., East and West Germany). The overall country averages were used in our calculations. We tested the correlations between 18 theoretically relevant constructs with GSTI and its four components (Table 3). While we primarily focus on validation of the overall index, attention is also devoted to the individual components since future researchers may want to focus on these separately when studying specific research topics.

As expected, the GSTI components correlated strongly with the theoretically relevant external validation criteria, with most correlations falling within the 99% statistically significant 0.40–0.80 range for the pair sample sizes of $n=35$ to 51 (Table 3). The composite GSTI correlates positively with broad range indicators including economic and human development, Schwartz's model values of egalitarianism and intellectual autonomy, political and economic freedom, competitiveness, and Transparency International's Corruption Perception Index. Gender tolerance correlates positively with gender egalitarian beliefs, as well as women's empowerment and global gender gap. Minority tolerance correlates negatively with power distance and societal hierarchy orientation to a lesser extent. Tolerance of immigrants correlates negatively with globalization practices and in-group collectivism. Finally, Religion correlates negatively with both Leung and Bond's (2004) religiosity index and Schwartz's embeddedness measure (a.k.a. conservatism).

In some cases, the correlations corresponding to the focal GSTI are not the largest. However, the internal reliability of the GSTI is very high, with the components closely intertwined, so it is not surprising to see that most external criteria correlated strongly with all four index components.

In order to further demonstrate the incremental contribution of our GSTI composite index over that of Das et al. (2008), we added the Das et al. correlation with our external comparison measures as the last column within Table 3. Our overall GSTI correlations are reasonably higher (marked with "c") for 13 of the external measures than those of Das et al. These external measures are used to validate the overall index, and the four GSTI components. The gender, immigration, and religion results are not surprising since they were overlooked within Das et al.'s index, but two of our three minority questions were included within the Das et al. 5-item index. With only one exception, the remaining external correlation pairs between these two composites are practically ties. Further, correlations comparing these two indexes with the annual Global Gender Gap Index for 2006–2010 revealed a GSTI correlation linearly trending upward over time from 0.81 to 0.84 (vs a practically constant 0.57–0.58 for the Das et al. index); while a meta-analysis of Hofstede's Power Distance data over the past three decades reveals a GSTI correlation of -0.77 vs -0.49 for Das et al.

Table 2 GSTI rank agreement and sensitivity with equally weighted benchmark rank

Country	Median of ranks	Bench mark equal weight rank	Median – Bench mark difference	Median rank 95% CI range	Sensitivity of median of ranks		
					Range	Range/ Median (%)	Majority within 5 ranks (%)
Sweden	1	1	0	0	1	100	100
Norway	2	2	0	0	1	50	100
Andorra	3	3	0	0	0	0	100
Switzerland	4	4	0	0	1	25	100
Germany	5	5	0	1	2	40	56
France	6	6	0	1	8	133	60
New Zealand	6	7	-1	1	3	50	86
The Netherlands	8	9	-1	0	2	25	100
Australia	9	10	-1	0	3	33	100
Finland	10	13	-3	1	3	30	53
Canada	11	11	0	0	9	82	78
Spain	12	8	4	1	10	83	72
Argentina	12	12	0	1	6	50	82
Slovenia	14	14	0	0	6	43	96
Italy	15	15	0	0	2	13	97
Britain	16	17	-1	1	7	44	93
United States	18	18	0	1	7	39	86
Taiwan	18	19	-1	0	8	44	96
Uruguay	19	16	3	0	8	42	74
Bulgaria	20	21	-1	0	5	25	69
Trinidad and Tobago	22	20	2	1	12	55	65
Cyprus	22	22	0	0	3	14	100
Serbia	24	24	0	0	5	21	87
Japan	24	25	-1	1	28	117	41
Brazil	25	23	2	2	10	40	57
South Korea	26	28	-2	2	19	73	48
Chile	27	26	1	1	6	22	73
Poland	28	29	-1	1	6	21	93
Moldova	29	32	-3	2	11	38	52
Peru	30	27	3	2	12	40	56
Romania	30	30	0	1	7	23	55
Ukraine	32	31	1	0	8	26	71
China	33	35	-2	1	6	18	90
Colombia	34	33	1	1	10	29	77
Mexico	34.5	34	0.5	1	6	17	96
Turkey	37	38	-1	1	5	14	94
Vietnam	37	40	-3	2	9	24	75
Ethiopia	38	36	2	1	5	13	92
Guatemala	40	37	3	3	13	33	45
South Africa	40	39	1	1	7	18	70
Rwanda	41	43	-2	3	8	20	51
Russia	42	42	0	0	4	10	88
Burkina Faso	43	41	2	1	5	12	95
Zambia	43	44	-1	1	5	12	93
India	45	45	0	0	7	16	85
Thailand	46	47	-1	0	3	6	100
Georgia	48	46	2	2	7	15	79
Indonesia	48	48	0	0	5	10	94
Iran	49	52	-3	3	7	14	60
Ghana	50.5	50	0.5	1	4	8	50
Morocco	51	49	2	3	9	18	54
Malaysia	52	51	1	0	6	12	62
Jordan	52	54	-2	0	8	15	84
Mali	54	53	1	0	5	9	100
Egypt	55	55	0	0	2	4	100
Iraq	56	56	0	0	1	2	100

Table 3 Criteria validation analysis

Measure	Composite	Gender	Minorities	Immigrant	Religion	Das et al.
<i>Composite Index</i>						
Economic Development, GDP ppp	0.78	0.72	0.46	0.73	0.76	0.76
Human Development Index	0.79^c	0.78	0.61	0.74	0.68	0.69
Egalitarian Commitment, Schwartz	0.70^c	0.82	0.49	0.52	0.59	0.63
Intellectual Autonomy, Schwartz	0.77^c	0.73	0.47	0.74	0.72	0.56
Civil Freedom	0.77^c	0.72	0.66	0.62	0.69	0.54
Political Freedom	0.68^c	0.64	0.58	0.57	0.62	0.46
Economic Freedom	0.69^c	0.66	0.53	0.58	0.61	0.60
Competitiveness Index (high = 1)	0.65	0.61	0.39	0.59	0.61	0.72^c
Corruption Perception Index (Corrupt = 0)	0.72	0.69	0.45	0.64	0.61	0.71
<i>Gender</i>						
Gender Egalitarianism, GLOBE values	0.80^c	0.84	0.75	0.65	0.69	0.62
Global Gender Gap	0.72^c	0.71	0.60	0.66	0.61	0.62
Women's Empowerment Index (GEM)	0.80	0.84	0.64	0.65	0.69	0.79
<i>Minorities</i>						
Power Distance, Hofstede	-0.80^c	-0.76	-0.68	-0.65	-0.61	-0.73
Hierarchy, Schwartz	-0.70^c	-0.70	-0.56	-0.53	-0.39	-0.45
<i>Immigration</i>						
Globalization KOF Index ^a	0.77^c	0.72	0.61	0.71	0.67	0.61
In-Group Collectivism, GLOBE values	0.76	0.74	0.56	0.70	0.66	0.76
<i>Religion</i>						
Religiosity, Leung and Bond	-0.85^c	-0.77	-0.60	-0.69	-0.84	-0.74
Embeddedness ^b , Schwartz	-0.88^c	-0.80	-0.63	-0.80	-0.80	-0.67

^aKOF Globalization Index source: <http://globalization.kof.ethz.ch>

^bAlso appears as "conservatism" in some publications (Schwartz 1994, 1999).

^cCases with the external validation item measures where GSTI composite or the Das et al. index correlation is meaningfully stronger than the other.

Notes: All correlations in this table are significant at the 95% level for our data sample sizes $N = 35-51$.

Approx. critical correlation values: 0.30 ($p < 0.05$); 0.38 ($p < 0.01$); 0.50 ($p < 0.001$).

Bold values indicate correlations significant at the 1% level, marked only for the GSTI composite index and one of its four corresponding dimensions' external validations.

Concluding Remarks

A nation's overall tolerance score may be a key to better understand its role in the global community (Das et al., 2008; Weldon, 2006). We add value to the tolerance literature by utilizing a large cross-country sample to develop an aggregate GSTI that incorporates four theory-based tolerance indicators, and expand it using multiple derivation methods of indexing and ranking, avoiding any weight assumptions and providing measures assessing each country's rank sensitivity or robustness. This provides a clearer picture of the global community that may facilitate future theoretical and empirical research. Understanding tolerance is increasingly important as we continue to break down national borders in our evermore-globalized environment.

Our overall contribution to the literature on social tolerance includes the following advancements. First, while a few prior studies have developed a single measure of tolerance (e.g., Das et al., 2008), they have failed to theoretically and empirically distinguish between different types of tolerance. Both theoretically and empirically, we demonstrate that different types of tolerance exist. Moreover, while they are empirically correlated, they are far from perfectly so, suggesting that they will have different predictive abilities in different circumstances.

Second, by developing an overall GSTI that can also be broken into its subcomponents, we help improve prior research focused on both overall tolerance and tolerance regarding specific issues. Regarding overall tolerance, prior measures have failed to capture a broad representation of tolerance, leading them to be biased towards the particular tolerance measures being examined. For example, Das et al.'s (2008) measure, while a noteworthy first step in this literature, includes a total of four equally weighted items in a single factor and does not address gender tolerance or immigration or religion issues. While, as noted above, our four components are correlated, they are far from near-perfectly correlated. These characteristics of our GSTI measure compared with prior tolerance measures suggest that prior studies focused on broad tolerance may have been incomplete in their tolerance specification by using less comprehensive measures, while studies focused on tolerance regarding a particular tolerance area may have not used the most specific measure possible. Furthermore, we demonstrated that our GSTI correlates reasonably higher with our external validation measures than Das et al.'s index – particularly for the overall composite index.

For IB researchers, the availability of a comprehensive measure of tolerance allows them to examine more confidently important questions regarding how a country's overall social tolerance level may relate to important country-level outcomes such as the ability to attract foreign direct investment or the ability to harness a country's full workforce potential. This measure may also help in better understanding the determinants of constructs such as country reputation (e.g., Newburry, 2012) and cultural intelligence (e.g., Thomas et al., 2015). Social tolerance also appears to be an important factor in cultural distance, a construct that has been shown to play a major role in various international business phenomena from entry modes to international collaboration (cf., Shenkar, 2001; Tihanyi, Griffith, & Russel, 2005). From a multinational company perspective, it allows the exploration of questions such as how operating in countries with different tolerance profiles impacts the coordination of MNC policies across countries. Additionally, the four individual dimensions within our measure provide tools for IB researchers to examine specific questions related to individual areas of tolerance. For example, how does gender-based tolerance or tolerance towards immigrants and minorities in a country impact the HR practices of MNC subsidiary managers towards these specific groups in a country, or even how ethics of executives are influenced by religion-based tolerance? Particularly given that our measures examine tolerance based on perceptions rather than actual practices, they provide insights into underlying viewpoints of a population, regardless of legal requirements, which may help MNC managers determine a truer account of tolerance in a country, and appropriate practices to address this.

Third, in contrast to Das et al. (2008) and all other prior tolerance studies, the major methodological contribution of this study is that we generated not one but 124 indices and ranks per country (after removal of outliers), the medians of which are taken as the consensus median index and the median of all these ranks for each country. In a sense, this approach may be viewed as using a repeated sampling. This allows two additional benefits: (a) weights driven by the data, not assumed equal or selected without sufficient justification by human "experts" (these calculated weights can assist experts in their deliberations for a consensus based on other criteria); and (b) four measures of variability of the ranks of each country to assess each country's rank sensitivity or robustness. Overall, while we recognize the contributions of prior studies, our procedure is much more compatible with the reliability and validity standards currently expected within top journals than prior measures.

Finally, limited meta-analysis evidence over several years suggests an upward trend over time for GSTI composite correlations with the external annual indexes of Global Gender Gap (while the Das et al. index remains practically constant around 0.58 six years after 2005), and with Hofstede's Power Distance over the past three decades (correlation -0.77 vs -0.49 for Das et al.).

The GSTI composite index presented herein reveals 99% significant correlations with all 18 external (sub)indexes examined within Table 3 at all five levels, with 15 of these correlations ranging from 0.70 to 0.88; for example, at the composite level with human and economic development, as well as corruption perception and to a lesser extent with social freedom (civil, political, and economic). We also examine correlates of each GSTI component, providing additional validation at this level – which is particularly strong for gender, immigration, and religion. It remains debatable, however, whether these relationships are causal.

The indexing and ranking methodology presented herein, or variations of it, could be applied to different datasets, provided that the data have been analyzed suitably and the constructs have been properly validated. Furthermore, our approach does not require prior benchmark indexes and ranks; that role can be played by setting one of the equally weighted results generated as the default benchmark. Since the data standardization and method combinations employed here are neither exhaustive nor the best of all possible ones, no indexing/ranking approach can be clearly superior in all aspects. What we advocate is that a systematic combination of multiple representative methods is a powerful tool to assess any published or calculated ranking that uses equal (assumed out of convenience) or unequal weights (often inadequately justified). Namely, it allows examining: (a) the extent to and conditions under which such prior single method rankings agree with multi-method median rankings, (b) any systematic biases of single method ranks (e.g., relative to external variables), and (c) how sensitive or robust each country rank is under different scenarios.

At a policy level, answers produced by such an “automated” multi-method system can only serve as a starting point to facilitate deliberations of experts, decision-makers, or policy officials. Involving them early in the “game” from the criteria phase is the key to successful implementation. In national rankings, equal weighted averages are commonly employed because they are not only simpler or the only way known, but also because they avoid public scrutiny or potential objections to differentiated weights. Even then, an approach like the one presented herein can assist decision-makers to understand and compare the reasons, magnitude, and impact of both procedures and new variables added. The joint report by the Organisation for Economic Co-operation and Development with the European Commission Joint Research Centre (OECD ECJRC, 2008) has advocated guidelines for composite indicators. However, such procedures or ones similar to those advanced in this study have not been widely adopted in management and international business journals.

Limitations and Future Research

While we believe our manuscript makes significant contributions, it is not without limitations that open up future research opportunities. First, our tolerance dimensions are not exhaustive and future research could examine other tolerance components. Second, individual-level tolerance research has considered both attitudes and behavioral intentions (e.g., Sagiv & Schwartz, 1995; Schmid, Hewstone, Tausch, Cairns, & Hughes, 2009), and future research could extend this distinction to the country level. Third, our tolerance component measures varied based on the

current state of acceptance of a group in society and their representation within the WVS. However, future research is needed to better understand these differences. Fourth, space restrictions prohibited including a detailed discussion of the variables used to validate the measures, and more generally, the convergent and divergent validity of the GSTI. While we contend that our tests produce a highly valid measure, future research could nonetheless develop these validations in further detail – in relation to the validation measures we examine along with other measures that should be theoretically similar or distinct, such as the related construct of acceptance. Additionally, newly identified cultural dimensions (e.g., Minkov & Hofstede, 2012) and/or approaches to conceptualizing and measuring culture (e.g., Caprar, Devinney, Kirkman, & Caligiuri, 2015) could be used for further validation. These might include relating social tolerance to culture at the transnational or subnational levels (e.g., Venaik & Midgley, 2015) or to culture conceptualized in terms of schemas or norms as opposed to the more commonly used values approach (e.g., Leung & Morris, 2015). Fifth, we chose variables that were theoretically close to our tolerance dimensions to validate our constructs. We leave the development of theory and testing regarding the impact of tolerance on other variables to future studies. Sixth, tolerance is only one component in predicting the actions of countries, and certainly cannot explain everything, just as no single variable can. Future research could examine the impact of tolerance vis-à-vis other country-level variables in predicting the actions of countries. Seventh, cross-cultural differences in survey response styles are likely present in WVS and remain partially unaddressed in our study. While we discussed earlier why adjusting the data for this did not seem appropriate for our study, and examined whether extreme response style and acquiescence bias (Marin et al., 1992) influence our results, this issue should be kept in mind when interpreting the results of the present study. We encourage future researchers to explore in more depth the effects of cross-cultural differences in response styles on the results of social tolerance national rankings.

While we include a sub-index based on religion in our overall GSTI, we realize there are limitations associated with this measure. First, items related to “God” and “churches/religious authorities” may have much stronger applicability within predominantly monotheistic religious cultures (e.g., Müller, 2009). Second, while religion has commonly been associated with intolerance towards women and homosexuals, as noted earlier, many religions subscribe to a version of the “Golden Rule” – “to do unto others as you would wish them to do unto you.” As such, care regarding a universal association of religion with intolerance should be taken, and future research might delve more deeply into this issue. Inclusion of this dimension is not meant to be taken as a condemnation of religion, which in its various forms is central to the lives of a large percentage of the world’s inhabitants, and certainly a source of benefits to many. Third, the religion component of our social tolerance index differs from other dimensions in that it takes a more broad-based institutional approach to tolerance. While this on one hand increases the scale’s scope and the religious dimension empirically fits overall well with the other dimensions, we nonetheless recognize this conceptual difference. Items within the religion dimension may be less reflective of actual tolerance and more indicative of a general condition that has been associated in prior literature with a lack of tolerance. This difference in approach should also be recognized.

The WVS has been criticized for some data design and collection issues that are also limitations inherited here. The methodological index and ranking development procedures described in this paper require complete data and are computationally very intensive, as is the missing value analysis (MVA) procedure used for imputation of missing values. Thus both are

not feasible to apply directly to big datasets like the 83,000 individual respondents in WVS. In this study, if no data imputation is desired or available, the respondent-level dataset without any missing values in any of the 13 items of each respondent will be reduced to 42,440 complete responses from 41 countries. This would have eliminated 15 countries, including eight major ones (Britain, Canada, China, France, Japan, the Netherlands, New Zealand, and Russia), which we deemed unacceptable as not representative and likely to induce larger standard errors due to smaller sample sizes.

Finally, several other indexing procedures for data standardization, weighting, and aggregation could be used here or by others, such as principal components and factor analysis scores, an array of utilitybased and multi-criteria decision aids, data envelopment or conjoint analysis, etc. We avoided using such procedures in this study since our focus has been to use only self-contained embedded analytical index approaches. Several methods could also be used with ranked data outside any indexing approach, such as Borda, Condorcet, Copeland, dominance, max agreement, min regret, or L1/L2 optimization norms (Zanakis, Antony, Nguyen, & Simpson, 2006).

In an era with increasing cross-border interactions (Buckley & Ghauri, 2004), tolerance is an important concern and may have far-reaching implications for business and public policy, particularly with respect to labor markets, personnel selection and management, diversity management in the workplace, and inter-group tension at the organizational and national levels. From a practical standpoint, the availability of an index of this type will help public policy officials to better understand issues related to tolerance in their home markets, and better predict the impact of policies regarding important segments of the population – such as immigrant employment issues, minority rights or religiosity's impact on politics and voting. For managers of multinationals, understanding tolerance issues in their host markets will help them better manage coordination of policies such as those related to managing diverse groups in the workforce across different subsidiary contexts. By developing a multi-factor social tolerance index, we have provided an aid to theory development and empirical studies not only regarding overall tolerance, but also related to four tolerance dimensions of particular importance – gender, minorities, immigrants, and religion. As such, we hope our study provides a valuable tool for future research development.

The data-driven methodologies employed in this study are directly applicable to the development of indexes and rankings within any application area of interest to researchers such as priority funding of projects, and rankings of colleges, journals, services or industries, and states or countries. In a business or political reality, such results can only be the starting point to facilitate the deliberations of decision-makers or policy officials in reaching an initial consensus on priorities for action. Since the data standardization and method combinations employed here are neither exhaustive nor the best of all possible ones, no indexing/ranking approach can be clearly superior to another in all aspects. What we advocate here is that a systematic combination of representative multiple such methods is a very powerful tool to assess any published or calculated rankings that so often use equal weights (assumed out of convenience) or unequal weights (usually inadequately justified).

Acknowledgements

The authors gratefully acknowledge the valuable suggestions and insightful extensive comments, provided by the Editor and the anonymous three reviewers, which led to substantial changes and

improvements in this paper. Any errors remaining are the responsibility of the authors. This work was partially supported by a research grant from the Florida International University Center for International Business Education and Research. The first author gratefully acknowledges the valuable assistance provided by his doctoral students, Gladys Simpson in the implementation of the methods described in this paper, and Maria Rickling and Barri Litt for the initial data and preliminary literature preparation. We also benefited from discussions on MVA with Dr. David Matheson, senior statistician of the IBM SPSS support team.

Note

¹ The median-based index and rankings are also closer to the commonly used equally weighted benchmark values than the mean-based ones, and thus potentially more “politically acceptable” to policymakers.

References

- Adler, N. J. 2002. Global managers: No longer men alone. *International Journal of Human Resource Management*, 13(5): 743–760.
- Au, K. Y., & Cheung, M. W. L. 2004. Intra-cultural variation and job autonomy in 42 countries. *Organization Studies*, 25(8): 1139–1362.
- Bardhan, K., & Klasen, S. 1999. UNDP’s gender-related indices: A critical review. *World Development*, 27(6): 985–1010.
- Buckley, P. J., & Ghauri, P. N. 2004. Globalization, economic geography and the strategy of multinational enterprises. *Journal of International Business Studies*, 35(2): 81–98.
- Bullough, A., Kroeck, K. G., Newbury, W., Kundu, S., & Lowe, K. B. 2012. Women’s political leadership participation around the world: An institutional analysis. *Leadership Quarterly*, 23(4): 398–411.
- Caprar, D. V., Devinney, T. M., Kirkman, B. L., & Caligiuri, P. 2015. Conceptualizing and measuring culture in international business and management: From challenges to potential solutions. *Journal of International Business Studies*, 46(9): 1011–1027.
- Cukur, C. S., de Guzman, M. R. T., & Carlo, G. 2014. Religiosity, values, and horizontal and vertical individualism-collectivism: A study of Turkey, the United States, and the Philippines. *Journal of Social Psychology*, 144(4): 613–634.
- Das, J., DiRienzo, C., & Tiemann, T. 2008. A global tolerance index. *Competitiveness Review: An International Business Journal*, 18(3): 192–205.
- DelCampo, R. G., Jacobson, K. J. L., Van Buren, H. J. III., & Blancero, D. M. 2011. Comparing immigrant and US born Hispanic business professionals. *Cross Cultural Management*, 18(3): 327–350.
- Dempster, A., Laird, N., & Rubin, D. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1): 1–38.
- Dijkstra, A. G. 2002. Revisiting UNDP’s GDI and GEM: Towards an alternative. *Social Indicators Research*, 57(3): 301–338.
- Dolan, S. L., Diez-Pinol, M., Fernandez-Alles, M., Martin-Prius, A., & Martinez-Fierro, S. 2004. Exploratory study of within-country differences in work and life values: The case of Spanish business students. *International Journal of Cross Cultural Management*, 4(2): 157–180.

- Ekici, T., & Yucel, D. 2015. What determines religious and racial prejudice in Europe? The effects of religiosity and trust. *Social Indicators Research*, 122(1): 105–133.
- Elsbach, K. D. 2005. Perceptual biases and misinterpretation of artifacts. In A. Rafaeli, & M. Pratt (Eds), *Artifacts and organizations: Beyond mere symbolism* 61–81. Mahwah, NJ: Lawrence Erlbaum.
- Foner, N., & Alba, R. 2008. Immigrant religion in the US and Western Europe: Bridge or barrier to inclusion? *International Migration Review*, 42(2): 360–392.
- Gerhart, B., & Fang, M. 2005. National culture and human resource management: Assumptions and evidence. *International Journal of Human Resource Management*, 16(6): 971–986.
- Hanges, P. J., & Dickson, M. W. 2004. The development and validation of the GLOBE culture and leadership scales. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta (Eds), *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.
- Hodges Persell, C., Green, A., & Gurevich, L. 2001. Civil society, economic distress, and social tolerance. *Sociological Forum*, 16(2): 203–230.
- Hofstede, G. 1980. *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.
- Hofstede, G. 2001. *Culture's consequences*, 2nd edn. Beverly Hills, CA: Sage. House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds) 2004. *Culture, leadership and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.
- Hwang, C. L., & Yoon, K. 1981. *Multiple attribute decision making: Methods and applications*. New York: Springer-Verlag.
- ILO. 2013. Labor migration. International labour organization, <http://www.ilo.org/global/topics/labour-migration/lang-en/index.htm>, accessed 4 August 2013.
- Jäckle, S., & Wenzelburger, G. 2015. Religion, religiosity, and the attitudes toward homosexuality – A multilevel analysis of 79 countries. *Journal of Homosexuality*, 62(2): 207–241.
- Knack, S., & Keefer, P. 1997. Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics*, 112(4): 1251–1288.
- Lenartowicz, T., & Roth, K. 2001. Does subculture within a country matter? A cross-cultural study of motivational domains and business performance in Brazil. *Journal of International Business Studies*, 32(2): 305–325.
- Leung, K., & Bond, M. H. 2004. Social axioms: A model for social beliefs in multicultural perspective. In M. P. Zanna (Ed), *Advances in experimental social psychology*. Vol. 36. 119–197. San Diego, CA: Elsevier Academic Press.
- Leung, K., & Morris, M.W. 2015. Values, schemas, and norms in the culture-behavior nexus: A situated dynamics framework. *Journal of International Business Studies*, 46(9): 1028–1050.
- Linstone, H. A., & Turoff, M. (Eds) 2002. *The Delphi method: Techniques and applications*, <http://is.njit.edu/pubs/delphibook/delphibook.pdf>, accessed 8 January 2016. Little, J. 1992. Regression with missing X's: A review. *Journal of the American Statistical Association*, 87(420): 1227–1237.
- Marin, G., Gamba, R. J., & Marin, B. V. 1992. Extreme response style and acquiescence among Hispanics. *Journal of Crosscultural Psychology*, 23(4): 498–509.

- Merriam-Webster Dictionary. 2010. Tolerance, <http://www.merriam-webster.com/dictionary/tolerance>, accessed 20 August 2010.
- Minkov, M., & Hofstede, G. 2012. Hofstede's fifth dimension: New evidence from the World Values Survey. *Journal of Crosscultural Psychology*, 43(1): 3–14.
- Müller, T. 2009. Religiosity and attitudes towards the involvement of religious leaders in politics: A multilevel-analysis of 55 societies. *World Values Research*, 2(1): 1–29.
- Muralidhar, K., & Zanakis, S. 1992. A simple minimum-bias percentile estimator for the location parameter of the gamma, Weibull & lognormal distributions. *Decision Sciences*, 23(4): 862–879.
- Newbury, W. 2012. Waving the flag: The influence of country of origin on corporate reputation. In M. Barnett, & T. Pollock (Eds), *Oxford handbook of corporate reputation* 240–259. Oxford: Oxford University Press.
- Newbury, W., Belkin, L. Y., & Ansari, P. 2008. Perceived career opportunities from globalization: Globalization capabilities and attitudes towards women in Iran and the US. *Journal of International Business Studies*, 39(5): 814–832.
- Newbury, W., Gardberg, N. A., & Sanchez, J. I. 2014. Employer attractiveness in Latin America: The association among foreignness, internationalization and talent recruitment. *Journal of International Management*, 20(3): 327–344.
- OECD & European Commission Joint Research Centre. 2008. *Handbook on constructing composite indicators: Methodology and user guide*. Paris: OECD.
- Parboteeah, K. P., Hoegl, M., & Cullen, J. B. 2008. Managers' gender role attitudes: A country institutional profile approach. *Journal of International Business Studies*, 39(5): 795–813.
- Pawlak, Z. 1991. *Rough sets: Theoretical aspects of reasoning about data*. Dordrecht: Kluwer.
- Permanyer, I. 2011. Accessing the robustness of composite indices rankings. *Review of Income and Wealth*, 57(2): 306–326.
- Pigott, T. D. 2001. A review of methods for missing data. *Educational Research and Evaluation*, 7(4): 353–383.
- Price, T. D., & Feinman, G. M. 1995. *Foundations of social inequality*. New York: Plenum.
- Rosenthal, R., & Rosnow, R.L. 1991. *Essentials of behavior research: Methods and data analysis*, 2nd edn. New York: McGraw-Hill.
- Sagiv, L., & Schwartz, S. H. 1995. Value priorities and readiness for out-group social contact. *Journal of Personality and Social Psychology*, 69(3): 437–448.
- Schmid, K., Hewstone, M., Tausch, N., Cairns, E., & Hughes, J. 2009. Antecedents and consequences of social identity complexity: Intergroup contact, distinctiveness threat, and outgroup attitudes. *Personality and Social Psychology Bulletin*, 35(8): 1085–1098.
- Schwartz, S. H. 1994. Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H. C. Triandis, C. Kagitcibasi, S. C. Choi, & G. Yoon (Eds), *Individualism and collectivism: Theory, methods and applications*, 85–119. London: Sage.
- Schwartz, S. H. 1999. A theory of cultural values and some implications for work. *Applied Psychology: An International Review*, 48(1): 23–47.
- Schwartz, S. H. 2008. Cultural value orientations: Nature and implications of national differences. Moscow: Publishing House of SU HSE, <http://blogs.helsinki.fi/valuesandmorality/files/2009/09/Schwartz-Monograph-Cultural-Value-Orientations.pdf>, accessed 8 January 2016.
- Seguino, S. 2011. Help or hindrance? Religion's impact on gender inequality in attitudes and outcomes. *World Development*, 39(8): 1308–1321.

- Shenkar, O. 2001. Cultural distance revisited: Towards a more rigorous conceptualization and measurement of cultural differences. *Journal of International Business Studies*, 32(3): 519–536.
- Simpson, G. 2011. A multi-method approach for the assessment of composite indices and rankings. Doctoral Dissertation, Florida International University.
- Smith, P. B. 2004. Acquiescent response bias as an aspect of cultural communications style. *Journal of Cross-cultural Psychology*, 35(1): 50–61.
- Steel, P., & Taras, V. 2010. Culture as a consequence: A multilevel multivariate meta-analysis of the effects of individual and country characteristics on work-related cultural values. *Journal of International Management*, 16(3): 211–233.
- Stone, D. L., Hosoda, M., Lukaszewski, K. M., & Phillips, T. N. 2008. Methodological problems associated with research on unfair discrimination against racial minorities. *Human Resource Management Review*, 18(4): 243–258.
- Thomas, D. C., Liao, Y., Aycan, Z., Cerdin, J. -L., Pekerti, A. A., Ravlin, E. C., Stahl, G. K., Lazarova, M. B., Fock, H., Arli, D., Moeller, M., Okimoto, T. G., & van de Vijver, F. 2015. Cultural intelligence: A theory-based, short form measure. *Journal of International Business Studies*, 46(9): 1099–1118.
- Tihanyi, L., Griffith, D. A., & Russel, C. J. 2005. The effect of cultural distance on entry mode choice, international diversification, and MNE performance: A meta-analysis. *Journal of International Business Studies*, 36(3): 270–283.
- Treisman, D. 2000. The causes of corruption: A cross-national study. *Journal of Public Economics*, 76(3): 399–457.
- Venaik, S., & Midgley, D. F. 2015. Mindscapes across landscapes: Archetypes of transnational and subnational culture. *Journal of International Business Studies*, 46(9): 1051–1079.
- WAIB. 2015. Creating and publishing high-quality research on women and gender. Panel sponsored by the Women of AIB (WAIB), Academy of International Business Annual Meeting, Bangalore, June.
- Weldon, S. 2006. The institutional context of tolerance for ethnic minorities: A comparative, multilevel analysis of Western Europe. *American Journal of Political Science*, 50(2): 331–349.
- Wrench, J. 2011. Data on discrimination in EU countries: Statistics, research and the drive for comparability. *Ethnic and Racial Studies*, 34(10): 1715–1730.
- WVS. 2005. World Value Survey, www.worldvaluessurvey.org, accessed 18 June 2010.
- Yeganeh, H. 2015. Religiosity, socio-economic development and work values: A cross-national study. *Journal of Management Development*, 34(5): 585–600.
- Zanakis, S. H., & Rickling, M. 2010. Global tolerance index and robust multi-method country rankings. CIBER Center Report June 01, Florida International University.
- Zanakis, S. H., Antony, S., Nguyen, V., & Simpson, G. 2006. Comparison of consensus ranking methods with an application to airlines service quality, MCDM Multi Criteria Decision Making 18th International Society Conference, Greece.

About the Authors

Stelios H Zanakis is Professor Emeritus of Decision Sciences & Information Systems, College of Business, Florida International University. His articles, mainly in multi-criteria decision support, multivariate data analytics, and knowledge discovery in databases, have been published

in Decision Sciences, Management Science, and European Journal of Operational Research, among others. He served as associate editor of the first and guest editor for the above other two journals. He received his PhD in Management Science from The Pennsylvania State University.

William Newburry is Chair of the Department of Management and International Business and the SunTrust Bank Professor at Florida International University. He received his PhD from New York University. His research focuses on how multinational corporations relate to subsidiaries and other local stakeholders when they invest overseas. He has published in top journals such as JIBS, Strategic Management Journal, and Organization Science, among others. He is the Series Editor of Research in Global Strategic Management, and Chair of AIB's Latin America Chapter.

Vasyl Taras received his PhD in Management from the University of Calgary, Canada and his Master's in Political Economy from the University of Texas at Dallas. He is an Assistant Professor of International Business at the University of North Carolina at Greensboro. He conducts research on cross-cultural and global virtual teams, crowd-based problem solving, and experiential learning. He is the founder of the X-Culture Project. He has lived, worked, and studied in half a dozen countries and has experience as a manager, entrepreneur, and business consultant.