Open access • Journal Article • DOI:10.1111/APPS.12095

Culture and Testing Practices: Is the World Flat? — Source link

Ann Marie Ryan, Matthew C. Reeder, Juliya Golubovich, James A. Grand ...+5 more authors

Institutions: Michigan State University, University of Maryland, College Park, University of Pretoria, Ghent University ...+2 more institutions

Published on: 01 Jul 2017 - Applied Psychology (Wiley)

Topics: Uncertainty avoidance and Personnel selection

Culture and Testing Practices: Is the World Flat?

Ann Marie Ryan*

Michigan State University, USA

Matthew C. Reeder

HumRRO, Alexandria, VA, USA

Juliya Golubovich

ETS, Princeton, NJ, USA

James Grand

University of Maryland, USA

Ilke Inceoglu

Surrey Business School, University of Surrey, UK

Dave Bartram

CEB's SHL Talent Measurement Solutions, and University of Pretoria, South Africa

Eva Derous

Ghent University, Belgium

Ioannis Nikolaou

Athens University of Economics and Business, Greece

Xiang Yao

School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, China

**Address for correspondence: Ann Marie Ryan, Michigan State University, 316 Physics Drive, East Lansing, MI 48824 USA. Email: ryanan@msu.edu*

RUNNING HEAD: Culture and testing

ABSTRACT

There has been much speculation regarding the influence of cultural norms on the acceptance and use of personnel selection testing. This study examined the cross-level direct effects of four societal cultural variables (performance orientation, future orientation, uncertainty avoidance, and tightness–looseness) on selection practices of organisations in 23 countries. A total of 1,153 HR professionals responded to a survey regarding testing practices in hiring contexts. Overall, little evidence of a connection between cultural practices and selection practices emerged. Implications of these findings for personnel selection and cross-cultural research as well as directions for future work in this area are described.

Keywords: selection, testing, culture, multi-country

INTRODUCTION

Over the past several decades, there has been a rise in multinational corporations and an accompanying rise in focus on using selection practices on a global basis (see Steiner, 2012 for a review). Best practice standards (e.g., International Test Commission, 2010; Ryan & Tippins, 2009) emphasize the importance of considering societal culture in the design and implementation of selection tools, including practices such as establishing equivalence of measures across cultures (Hambleton, 2005; Meade & Lautenschlager, 2004), conducting reviews of test content for cultural sensitivity (Brislin, 1986), and assessing applicant reactions cross-culturally (Ryan et al., 2009; Steiner & Gilliland, 2001). Specifically, much has been written about how societal culture may influence the acceptance and use of testing (Fell, Konig & Kammerhauf, 2015; Fell & Konig, 2016: Lim, Chavan & Chan, 2014; Ryan et al., 2009; Ryan, McFarland, Baron & Page,

1999; Sandal et al., 2014: Steiner & Gilliland, 1996). This issue of where there is societal cultural variability in selection tool use and acceptance is particularly important from a practical side, as it can affect whether resources are devoted to developing culturally-specific assessments and policies versus using more globally-standard approaches, as well as heighten or reduce concerns regarding practice effectiveness in different locations.

Despite this widespread emphasis on societal cultural differences in the acceptability of selection practices, the few existing large scale studies do not show strong support for an influence of societal culture on perceptions of selection method utility (Ryan et al., 1999; Ryan et al., 2009). To be sure, there are certainly *country* differences in the use of selection methods as economic, educational, and legal distinctions across countries can influence the ability and desire to use certain testing tools (Ryan & Tippins, 2009). Also, specific assessments may or may not exhibit measurement invariance across cultures (e.g., Bartram, 2013; Dai, Han, Hu & Colarelli, 2010) and scores may require culturally-based normative interpretations (Bartram, 2008; Fell & Konig, 2016). However, the role of *societal culture* as an influence on selection practices is less well-investigated. Indeed, the paucity of research available to guide the selection practitioner faced with expanding systems across borders (Ryan & Tippins, 2009) may lead to a slower spread of selection innovations, unnecessary restrictions in adopting methods, and general reluctance to take on the daunting task of justifying selection method use country by country. Ryan, Wiechmann, and Hemingway (2003) urged HR practitioners to investigate the legitimacy of pressures to differentiate hiring processes on the basis of culture rather than assuming modifications to selection processes are needed.

This study contributes to research on global applications of assessment by investigating the potential role of societal cultural differences. The most comprehensive investigation of the

role of societal culture in selection practice use, Ryan et al. (1999), is limited because: a) it predates much of the globalization of work that exists today; b) much has changed in assessment, particularly with regard to technology; c) it focused on Hofstede's cultural values framework (1980) but since that time the GLOBE framework (House, Hanges, Javidan, Dorfman & Gupta, 2004) has emerged as an empirically-based approach to examining societal culture; d) additional rigorous multi-country studies have further revealed promising new perspectives on the cultural difference of tightness-looseness (Gelfand et al., 2011); and e) the development of hierarchical linear modeling methods enables an examination of culture's influence on selection practices that is more methodologically appropriate than earlier analytic approaches for considering nesting of organizations within countries (Gelfand, Leslie, & Fehr, 2008). To address these theoretical and methodological limitations, we examined cross-level direct effects of cultural characteristics (at the societal level) on selection practices (at the organizational level). We first provide a brief review of previous work examining the relationship between societal culture and personnel selection testing. We then propose how four cultural variables may influence testing practices in employment contexts, and describe a 23-country study that evaluates those hypotheses.

Culture and Selection Practice: Literature Review

When using a selection tool globally, one must deal with all the typical challenges of designing and implementing a selection system for large scale implementation. Additionally, international selection contexts require further considerations regarding translation across languages and differences in legal constraints, labor forces, technology, and cultural acceptability (Ryan & Tippins, 2009; Steiner, 2012). Numerous surveys have revealed differences in the use and comparability of selection practices between countries (e.g., Clark, 1993; Huo, Huang &

Napier, 2002; Krause & Thornton, 2009; Salgado & Anderson, 2001; Shackleton & Newell, 1994). There have also been examinations of differences in the perceptions of selection procedures across countries (e.g., Anderson, Salgado & Hulsheger, 2010; Steiner & Gilliland, 1996; Steiner & Gilliland, 2001). However, such studies have not empirically addressed whether societal cultural characteristics might account for this variability rather than labor markets, legal constraints, educational systems, or the like which vary across countries. As Ryan et al. (1999) noted, understanding *why* differences in selection practices exist is particularly important for the multinational HR practitioner, as it provides guidance as to how differences might be addressed, accommodated, or, in some cases, suggest a strong obstacle to globalization of a practice.

A small number of notable studies have sought to more directly examine the specific role of societal cultural characteristics in selection practices. For example, Ryan et al. (1999) surveyed the selection practices of 959 organizations across 20 countries. They found that companies in nations valuing uncertainty avoidance (i.e., seeking to reduce the unpredictability of the future, as indexed by Hofstede, 1980) tended to use more tests, use them more extensively, and audit their processes to a greater extent, but they also used less varied methods of selection. Nevertheless, and despite revealing considerable variability in selection practices across nations, societal cultural values accounted for little of the between-country variance in selection practices. At the time, Ryan et al. noted "If the differences are due to lack of resources or the spread of technology, the rise of multinationals and the general trend toward globalization should lead to their dissipation" (p. 362). Given that the intervening years since that study have seen those changes, one might expect less variability across nations in testing practices today.

In 2009, Ryan et al. surveyed 1,199 individuals in 21 countries about their individual cultural values (i.e., personal endorsement of a value) and their perceptions of eight selection

methods. Respondents' perceptions of biodata, personality inventories, and cognitive ability tests were weakly related to the achievement/ascription values of individuals (beliefs that social status, titles, and the like are important to consider versus personal accomplishments). Overall however, cultural values were once again minimally and inconsistently related to differences in perceptions of selection methods, leading the authors to conclude that individuals' perceptions of selection methods were more similar than different globally. Notably though, Ryan et al. (2009) found that gross domestic product (GDP) was predictive of differences in respondents' fairness perceptions across countries. That is, those in nations with little employment opportunity viewed testing more positively, likely because it indicated a more level playing field with regard to economic advancement.

In the years since these studies, much has changed in both the global landscape of selection practices as well as current conceptualizations of societal culture. With respect to the former, technological advances in personnel assessment and selection delivery methods have made computerized and internet-based testing technologies more accessible and prevalent, raising many questions regarding validity, reliability, and test security (see Tippins, 2015 for a review). With respect to the latter, developments in the study and understanding of cultural characteristics have emerged that may reveal new and important insights. For example, the GLOBE research program (see House, Hanges, Javidan, Dorfman, & Gupta, 2004 for an overview) has emerged as one of the most comprehensive and widely cited works on societal cultural characteristics. Utilizing data from over 17,000 managers in 951 organizations across 62 counties, the GLOBE program empirically established nine distinct cultural dimensions characterizing the norms, practices, and beliefs of a society. Additionally, Gelfand and colleagues' (2011) recent work has found considerable support for the degree to which deviance

from societal norms and expectations is permitted (i.e., tightness-looseness) to be a key differentiator among cultures.

In sum, there are a number of studies documenting differences in selection practices and perceptions of selection methods across countries, but these tend to show considerable cross-country similarity on the whole (Steiner, 2012). When country-level differences have been investigated, societal cultural characteristics have not been found to be strong correlates of those differences. More recently, Fell and Konig (2016) did find connections of cultural characteristics to engaging in faking behavior, although not in the ways anticipated. The most directly applicable examination of societal cultural values in relation to practices by Ryan et al. (1999) is largely outdated in terms of the state of practice examined, its basis for defining and assigning cultural characteristic to nations, and the analytic approaches used to evaluate multilevel hypotheses. The globalization of business provides a much different landscape for organizations now relative to the 1990s; hence, there is a need for a new examination of societal cultural characteristics and selection practices.

Consequently, we focused on four characteristics identified by contemporary treatments of societal culture as potentially relevant to test use in selection contexts: **(1)** performance orientation, **(2)** future orientation, **(3)** uncertainty avoidance, and **(4)** tightness-looseness. Note that while Ryan et al.'s (1999) examination included uncertainty avoidance, the other three cultural characteristics have not been examined in prior research on testing practices. Focusing on assessment practices, rather than selection more broadly, allowed us to achieve a desirable level of specificity while keeping our survey at a manageable length. In terms of specific selection practices, we primarily examined practices related to testing that could be more objectively reported (e.g., general types of testing used, use of security measures) and that are of

most interest to those engaged in employment testing in large organizations and consulting contexts (Farr & Tippins, 2010).

Cultural Practices and Test Use: Hypotheses

Performance Orientation has been defined as the degree to which a culture encourages and rewards excellence and improvement in performance (Javidan, 2004). Highly performance-oriented societies value competitiveness, reward individual achievement, and emphasize results, suggesting a willingness to use assessments to evaluate individuals so as to determine who is most deserving. Further, some selection tools (e.g., work samples) are specifically measures of performance, and hence may have greater usage in highly performance-oriented cultures. Finally, Javidan (2004) also notes that highly performance-oriented societies see formal performance feedback as necessary for facilitating improvement. In a number of countries in Western Europe, for example, feedback to applicants is an expected, obligatory component of the hiring process (Bartram, 2001; Schinkel, van Dierendonck, van Vianen, & Ryan, 2011). In cultures where performance orientation is lower, salient evaluation and formal feedback in assessment contexts are likely to be seen as discomforting and less expected. These considerations led us to propose that:

> H1a: Organizations in highly performance-oriented cultures will be more likely to use tests than those in low performance-oriented cultures.
>
> H1b. Organizations in highly performance-oriented cultures will be more likely to use procedures that elicit performance-relevant behaviors and skills than those in low performance-oriented cultures.

H1c: Organizations in highly performance-oriented cultures will be more likely to give applicants feedback than those in low performance-oriented cultures.

Future Orientation has been defined as the extent to which a culture encourages delay of gratification and rewards investment in the future (Ashkanasy, Gupta, Mayfield, & Trevor-Roberts, 2004). Selection is inherently a process of predicting future behavior, so organizations in future-oriented societies may be more likely to use personnel testing for selection, see investment in selection assessments as useful, and invest in a greater range of assessment tools. Finally, because there is a greater willingness to defer gratification and to take a longer-term view in future-oriented cultures (Ashkanasy et al., 2004), there may also be a greater willingness to use assessment tools that require greater investment in development/administration (i.e., customized rather than off-the-shelf). Consequently, we expected that:

H2a: Organizations in highly future-oriented cultures will be more likely to use testing than those in low future-oriented cultures.

H2b: Organizations in highly future-oriented cultures will be more likely to use a greater variety of assessment techniques than those in low future-oriented cultures.

H2c: Organizations in highly future-oriented cultures will be more likely to use more customized/self-developed tools than those in low future-oriented cultures.

Uncertainty Avoidance is the extent to which ambiguous situations are seen as threatening and a society relies on social norms and rules to alleviate the unpredictability of future events (Sully de Luque & Javidan, 2004). Javidan and House (2001) note that high

uncertainty avoidance translates into a desire to have rules, procedures, and structures to manage daily situations, suggesting that organizations in high uncertainty avoidance societies would desire standardized selection processes and the use of assessments as a means of reducing ambiguity regarding the hiring process. Indeed, assessment itself may be viewed as an "uncertainty-reducing technology" (p. 607, Sully de Luque & Javidan, 2004). When there is a greater tolerance for ambiguity (low uncertainty avoidance), there also may be less of a concern about whether or not an individual test taker has cheated, and hence a lower emphasis on proctoring, monitoring, and other security-related practices. Indeed, "less tolerance for breaking rules" is listed as a core characteristic of high uncertainty avoidance societies (p. 618, Sully de Luque & Javidan, 2004). Therefore, we expected that:

H3a: Organizations in high uncertainty avoidance cultures will be more likely to use a variety of assessment techniques than those in low uncertainty avoidance cultures.

H3b: Organizations in high uncertainty avoidance cultures will be more likely to have greater standardization in selection processes than those in low uncertainty avoidance cultures.

H3c: Organizations in high uncertainty avoidance cultures will be more likely to use more test security methods than those in low uncertainty avoidance cultures.

H3d: Organizations in high uncertainty avoidance cultures will be less likely to use unproctored testing than those in low uncertainty avoidance cultures.

Tightness-looseness reflects the strength of social norms and whether there is punishment for violations. Organizations in tight societies are proposed to emphasize predictability. Gelfand,

Nishii, and Raver (2006) explicitly suggested a connection of this construct to selection: "We expect that selection and recruitment strategies are much stronger in organizations in tight versus loose societies, meaning that organizations in tight societies seek to restrict the range of individuals who enter the organization and to select individuals who match the organizational culture, to a greater extent than do organizations in loose societies" (p. 1232). Gelfand et al. (2006) further note that increasing the reliability of information about potential employees is an important aim for recruitment and selection strategies in tight societies, suggesting test use might be greater in those locations.

Gelfand et al. (2006) also suggest that the types of tests implemented by organizations in tight versus loose societies are likely to differ. In loose societies, there may be more variability in individual applicants' experiences as there are likely to be fewer strong, salient societal norms concerning what constitutes important developmental opportunities, educational experiences, etc. As a result, individual difference measures of knowledge, skill, and ability may be of particular value in loose societies for distinguishing more qualified from less qualified workers. Alternatively, tight societies are likely to exhibit less variance across individuals in their backgrounds, experiences, etc. as norms about these activities are more highly prescribed. Consequently, the match of the person to the organization may be of particular importance in tight societies where deviance from norms is less tolerated. As such, tight societies may be more likely to focus on value and personality fit attributes in their selection assessment practices to identify desirable applicants, even though societal conformity pressures may mean less applicant variability in responses. Further, the emphasis on rules and conformity will relate to less tolerance of cheating, so we would expect organizations in tight societies to place a greater emphasis on security procedures and test supervision. Based on this logic, we hypothesized:

H4a: Organizations in tight societies will be more likely to use testing than those in loose societies.

H4b: Organizations in tight cultures are less likely to select on constructs within the domains of knowledge, skill, and ability than those in loose cultures.

H4c: Organizations in tight cultures are more likely to assess personality, work styles, and other personal characteristics than those in loose cultures.

H4d: Organizations in tight societies will be more likely to use more test security methods than those in loose societies.

H4e: Organizations in tight societies will be less likely to use unproctored testing than those in loose societies.

In sum, our study draws on contemporary research on culture which suggests that the shared norms, ideals, and principles of the societies in which modern organizations operate may influence a company's selection practices. Cross-cultural theories and research posit that such commonly held beliefs should be observable as "top-down" effects that both shape and constrain the likely behaviors expressed by individuals within a society (Gelfand et al., 2008). We thus focus specifically on those societal cultural characteristics which theory suggests are likely to influence the beliefs and perceptions of organizational decision-makers regarding human resource practices related to testing and assessment during hiring.

METHOD

Sample

One limitation of Ryan et al.'s (1999) study was under-sampling and low response rates in certain countries. In some ways, this is inevitable, as markets differ by region and finding HR personnel in some regions is a much more challenging task than in others. Given that the goal of the present study was to evaluate cultural- rather than country-level differences, our sampling strategy focused on the societal cultural dimensions of interest rather than maximizing representation of multiple countries. Thus, we sought to ensure that we would have organizational participants from countries high and low on the specific cultural characteristics of interest, rather than focusing on a certain number of countries or a certain volume of participants within a country. To do so, we relied on the cultural score bands provided by the GLOBE research program (House et al., 2004) for performance orientation, future orientation, and uncertainty avoidance to identify potential countries from which to sample. The cultural score bands cluster countries according to relative similarities in reported cultural norms; thus, countries from within the same band for a given societal norm/value (e.g., performance orientation) share similar—but not identical—perceptions of that societal norm/value. Consequently, our sampling strategy involved purposefully attempting to sample organizations from multiple countries that spanned the cultural score bands so as to ensure reasonable variance in cultural practices, while also taking into account logistical considerations concerning translation needs and access to HR professionals in those locations. On the basis of these desiderata, we initially identified 28 countries on which to focus our sampling efforts. However, we were unable to obtain responses from organizations in four of these countries (Dubai, Egypt, Mexico, and Zimbabwe). Additionally, responses from Saudi Arabia could not be included in

any of the hypothesis analyses as no cultural scores were available for this country nor could they be imputed (though responses from Saudi Arabia are included in descriptive statistics for completeness). Thus, our final sample for hypothesis testing included organizations from 23 different countries that varied in their cultural practices (Table 1).

Table 1. Cultural Practice Values by Country

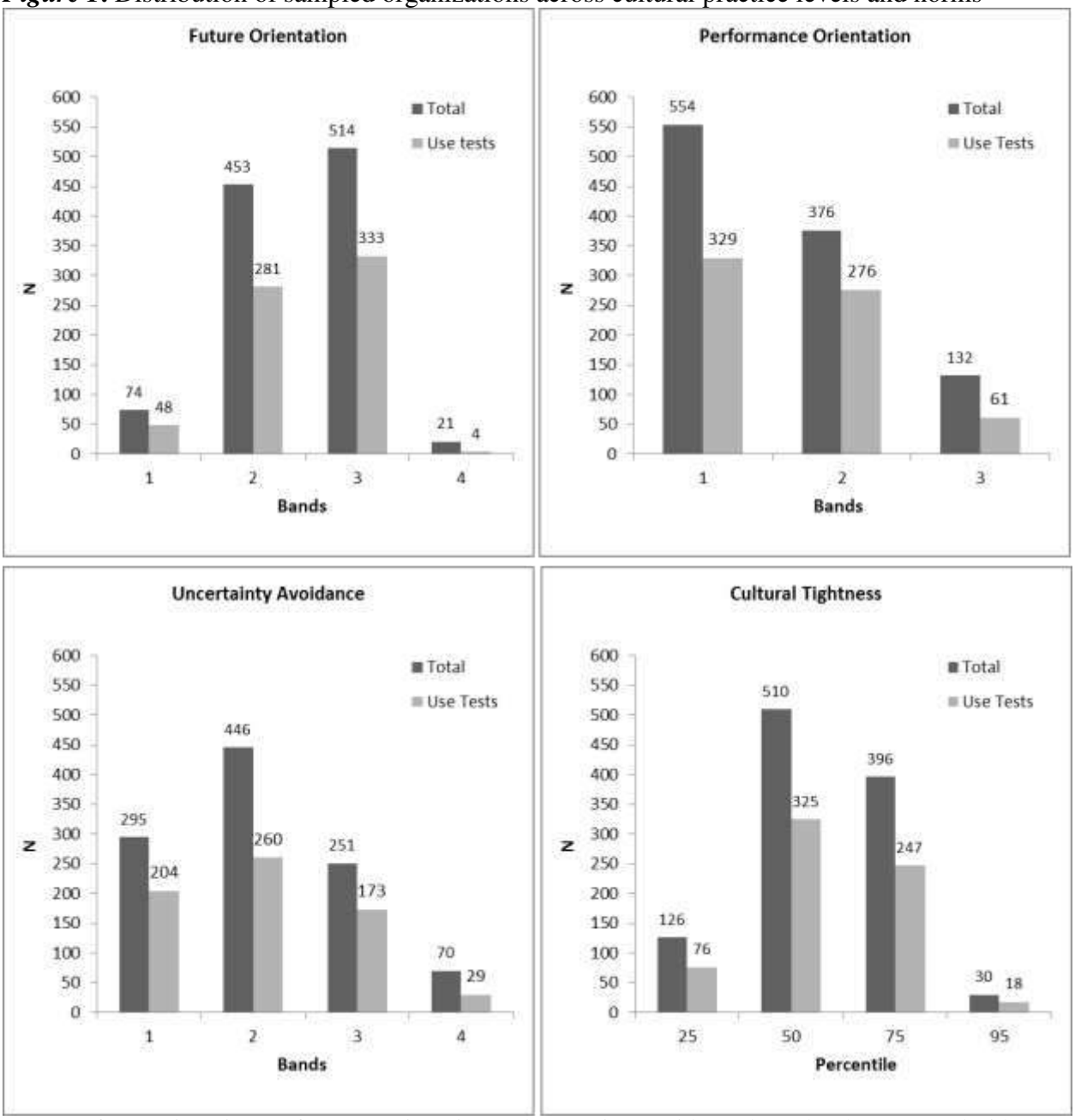
	Performance Orientation	Future Orientation	Uncertainty Avoidance	Cultural Tightness	Number of Organizations
Australia	4.37	4.09	4.40	4.40	20
Belgium	3.81[a]	3.61[a]	3.84[a]	5.60	181
Brazil	4.11	3.90	3.74	3.50	3
China	4.37	3.68	4.81	7.90	180
Denmark	4.40	4.59	5.32	7.12[a]	1
France	4.43	3.74	4.66	6.30	34
Germany	4.42	4.41	5.35	6.50	19
Greece	3.34	3.53	3.52	3.90	49
Hong Kong	4.69	3.88	4.17	6.30	14
India	4.11	4.04	4.02	11.00	20
Indonesia	4.14	3.61	3.92	9.44[a]	4
Italy	3.66	3.34	3.85	6.80	22
Netherlands	4.46	4.72	4.81	3.30	71
New Zealand	4.86	3.46	4.86	3.90	3
Portugal	3.65	3.77	3.96	7.80	40
Russia	3.53	3.06	3.09	7.49[a]	21
Saudi Arabia[b]	--	--	--	--	2
Singapore	4.81	4.88	5.16	10.40	2
South Africa	4.01	4.08	4.06	6.49[a]	1
Spain	4.00	3.52	3.95	5.40	1
Sweden	3.67	4.37	5.36	7.51[a]	90
Turkey	3.82	3.74	3.67	9.20	4
United Kingdom	4.16	4.31	4.70	6.90	22
United States	4.45	4.13	4.15	5.10	260

Note. Cultural practice values for Performance Orientation, Future Orientation, and Uncertainty Avoidance are from Table B.2 in Hanges (2004a). Cultural practice values were computed using a data transformation procedure designed to correct for potential societal-level response bias. Cultural tightness values are from Table 1 in Gelfand et al. (2011).

[a]Values imputed based on the Hofstede dimensions.

[b]Values for country not available and could not be imputed; consequently, responses from this country were excluded from subsequent analyses

Figure 1. Distribution of sampled organizations across cultural practice levels and norms



Note. For future orientation, performance orientation, and uncertainty avoidance, lower bands correspond with higher levels of a cultural practice. Only three cultural score bands were provided for performance orientation by Javidan (2004). For cultural tightness, higher percentiles correspond with higher levels of cultural tightness.

Figure 1 provides a graphical summary of the total number of organizations obtained from each of the cultural score bands for future orientation, performance orientation, and uncertainty avoidance in the final sample, as well as the number of organizations that reported

using tests as part of their hiring practices within each band. Given that no comparable bands were available for the cultural tightness variable, Figure 1 provides a summary for cultural tightness in which organizations are clustered based on percentile ranks on cultural tightness relative to the data reported in Gelfand et al. (2011). Overall, these graphs suggest that reasonable variance in societal cultural characteristics was obtained across the organizations included in the final sample.

Between March 2011 and March 2012, we targeted HR professionals by contacting professional associations related to testing and selection, LinkedIn groups for HR professionals, and individuals on a marketing e-mail list of a major test publisher in each country. It was thus impossible to calculate an accurate response rate as the true population of HR professionals with internal responsibilities for selection systems was unknown. A total of 1,153 HR professionals participated in the online questionnaire and indicated a specific country (versus "other") as their company's home office. After removing 89 cases (7.7%) from respondents whose company was already represented in the dataset[1], the largest representation in the sample was from the U.S. (24.4%), Belgium (17%), and China (16.9%). Listed in order of representation, other countries included Sweden, the Netherlands, Greece, Portugal, France, the United Kingdom, Italy, Russia, Australia, India, Germany, Hong Kong, Indonesia, Turkey, Brazil, New Zealand, Saudi Arabia, Singapore, Spain, Denmark, and South Africa. The industry sectors associated with the largest representation were professional services (19.2%), manufacturing (17.5%), financial (8.4%), retail (6.7%), health care (6.5%), telecommunications (3.6%), and transportation (3.1%).

[1] In cases of multiple respondents per company, either within a single country or across different countries, we randomly chose one respondent from the company to keep for analyses. We could only identify "duplicate" respondents to remove in cases where respondents provided an interpretable company name (63% of all respondents).

Respondents were primarily HR managers (28.6%) and HR executives such as director or vice president (25.4%).

Survey Content

The survey content covered decisions to use tests, descriptions of testing programs for those who did use tests, and testing policies and practices[2]. Individuals were asked whether they used tests for entry-level management employees, as testing practices can vary widely by level. The survey was developed by the authors with review from several other testing and assessment experts to capture a wide range of potential practices (e.g., hurdles, use of adaptive testing); however, we limit our focus here to assessment practices expected to vary based on culture. The specific items related to our hypothesis tests are described in Table 2.

Table 2. Survey Items

Item	Response Scale	Hypotheses
Do you use tests for selecting entry-level management employees as part of your typical hiring process?	Yes/no	1a, 2a and 4a
Which of the following characteristics do you assess with your tests? (check all that apply)	Abilities (e.g., math, verbal, language), personality (e.g. conscientiousness, adaptability, work styles), experience (e.g., background), knowledge (e.g., job specific technical knowledge), social skills (e.g., interpersonal skill, social perceptiveness), administrative skills (e.g., planning, organizing), leadership competencies, interest, work values (e.g., autonomy), motivation (e.g., achievement motivation), other. Responses yes/no to each question; also summed number of characteristics assessed	1b, 2b, 3a,4b
How often is feedback provided to applicants on their test results?	1-5; never to almost always	1c
Which of your paper-and-pencil tests do you administer in unsupervised settings? Which of your computerized tests do you administer in unsupervised settings?	List of tests from above scored as yes, use unsupervised; none scored as no	3d and 4d

[2] The survey questionnaire is available upon request.

How frequently have you disqualified applicants for cheating on your paper-and-pencil tests?	1-5; never to very frequently	3c and 4c
Which of the following best describes how the tests used in your selection process were created? (Check all that apply)	a)Tests were created solely by individuals in our organization; b) Tests were created solely by individuals external to our organization (i.e., purchased from/developed by consultant, professional test publisher, etc.); c) Tests were created through collaboration between individuals in our organization and tests were created by individuals external to our organization. Scored as a and c= customized and b= off the shelf	2c
Do you use any of the following security measures when administering paper-and-pencil tests? Do you use any of the following security measures when administering supervised computerized tests? Do you use any of the following security measures when administering unsupervised computerized tests?	List of measures depending on type of testing environment (e.g., store test materials in locked area; use of keystroke analyses); scored as 0 = none 1= any	3c and 4c
How frequently have you disqualified applicants for cheating on your supervised (unsupervised) computerized tests?	1-5; never to very frequently	3c and 4c
Many countries have data protection and privacy laws that are important to consider when testing. In your organization's testing process, do you and/or your test vendor	List of practices (e.g., use firewalls and password protection, have disaster recovery plans in place) Yes/no to each option; also scored total indicated as 0= no methods and 1= use of any method	3c and 4c
How standardized are your multinational testing practices	Same for all countries,, different across countries, mixture of custom and standardized	3b

In order to ensure adequate translation of survey items, we employed professional translators with a back translation check. Because assessment and HR practices can be jargon-laden (e.g., unproctored internet testing), we sought assistance from selection experts with facility in the languages of focus as well to ensure translation adequacy.

Cultural Practices

To obtain cultural scores for performance orientation, future orientation, and uncertainty avoidance for the countries included in our sample, we used the cultural practice scores reported

by the GLOBE research program corrected for response bias (see Hanges, 2004a, Table B.2, pp. 742-744). In the GLOBE research program, researchers collected data on cultural beliefs regarding whether a given value reflected the society "as it seeks to be" (cultural values) or reflected the society "as it currently is" (cultural practices). Given that the goal of the present study was to examine the degree to which the actual experience/expression of certain societal cultural norms influence an organization's selection assessment practices, we elected to use the cultural practice rather than cultural values scores to represent the societal cultural milieu in which organizations operate (see Atwater, Wang, Smither, & Fleenor, 2009; Ott-Holland, Huang, Ryan, Elizondo, & Wadlington, 2013, for similar views). To obtain data for the tightness-looseness of a society, the values reported by Gelfand and colleagues (2011) for the countries in our sample were used. Countries' scores on these four cultural characteristics are reported in Table 1. For a small number of sampled countries, complete data for either the GLOBE cultural dimensions or tightness/looseness were not available. To maximize the sample size available for hypothesis testing, we attempted to leverage previous data based on Hofstede's (1980) cultural dimensions to impute the necessary GLOBE or tightness/looseness values using a regression-based approach[3]. As shown in Table 1, only eight cultural practice scores across six countries required such imputation. This represented a relatively small proportion (8.7%) of the cultural practice scores used in our analyses.

Analytic Approach

Respondents were sampled from organizations across the world; as a result, organizations generally can be considered as nested within countries[4]. Consequently, our data were structured

[3] Additional information on this procedure can be obtained from the first author.
[4] Note this is a somewhat imprecise categorization as individuals responded based on their home office (i.e., where they were located). Some organizations (31.4%) reported testing globally but fewer of these used standardized practices so reported practices likely reflect the locale rather than a multinational policy.

such that multiple organizations were assumed to be sampled from each country, and each organization was associated with only one country. The societal culture predictor variables exist at the country level, whereas the testing practice and policy outcomes exist at the level of the individual organization. Given this design, analyses were carried out within a multilevel modeling framework, with organizations treated as level-1 units nested within countries, which were treated as level-2 units:

$$_{jk} = \pi_{0k} + r_{jk} \quad (1)$$

$$\pi_{0k} = b_{00} + b_{01} * X_k + u_{jk} \quad (2)$$

In the equations above, *j* indexes the j^{th} organization nested within the k^{th} country. The level-1 equation, Eq. 1, describes the testing policy and practice outcomes at the organizational level. The level-2 equation, Eq. 2, relates testing policy and practice outcomes to country-level predictors. The outcome, *y*, is modeled at the organizational level (level-1) as having a component that is common to all organizations within that country (π_{0k}) plus a component unique to each organization, r_{jk}. Variability in π_{0k} across countries is modeled using country-level predictors (i.e., cultural practices), X_k, the effects of which are captured by the slope coefficients in Eq. 2 (e.g., b_{01}). Support for the study hypotheses is evidenced by significant slope terms associated with the hypothesized predictors in the level-2 equations.

All hypotheses were tested using mixed models estimated with the *lme4* package 1.1-7 (Bates, Maechler, Bolker, & Walker, 2013) in R 3.1.1 (R Core Team, 2014). Linear or logistic models were fit as appropriate to the outcome in question. Outcomes were modeled with either normal or binomial distributions as appropriate using the standard linear mixed and logistic mixed models, respectively. Each outcome was regressed on all four cultural practice predictors

concurrently. Predictors were grand-mean centered prior to entry into the models. Analyses pertaining to whether or not organizations used testing at all utilized the entire sample. All other analyses included only organizations that indicated they conduct testing.

Note that although one can use these data to describe practices by country or by region, we do not do so here for several reasons. First, as our sampling strategy was to obtain organizations with a range of cultural characteristics at the country level of analysis, we have small numbers of organizations for many countries, which would provide an imprecise view of practices in those countries. Second, the GLOBE study identified 10 cultural clusters of countries with similar patterns of cultural characteristics (Javidan, House, & Dorfman, 2004). Although we do have representation for 9 cultural clusters, the sample sizes in two of those (Eastern Europe and Southern Asia) are quite small. Third, we did not hypothesize that a cluster of countries along with the concomitant economic and political differences would be a source of testing practice differences, but derived our hypotheses directly based on societal cultural characteristics. Finally, a descriptive summary of testing practices globally (not by country or cluster) is reported elsewhere (Ryan et al., 2015).

RESULTS

Table 3 provides descriptive statistics for the outcome variables used for testing Hypotheses 1-4. Overall, selection tests were used by 62.8% (n = 668) of the organizations sampled. Of those organizations that used tests for hiring purposes, 57% reported using some form of customized, as opposed to off-the-shelf/externally developed, assessment tools. Personality (84% of test-using organizations), abilities (80%), and leadership competencies (64%) were identified as the most commonly assessed construct domains. Organizations which used testing also reported assessing an average of five different characteristics as part of their

Table 3. Descriptive Statistics for Outcome Variables

	Mean	*SD*	*Min*	*Max*	*n*
Characteristics of Test Administration					
Use tests for selecting entry-level management employees[a]	0.63	0.48	0	1	1064
Use customized testing procedures[a]	0.57	0.50	0	1	668
Use unsupervised paper & pencil tests[a]	0.58	0.49	0	1	668
Use unsupervised computerized tests[a]	0.76	0.43	0	1	668
Frequency of feedback to applicants[b]	3.80	1.42	1	5	657
Standardization of multinational testing practices[c]	2.00	0.90	1	3	202
Characteristics Assessed with Tests[a]					
Ability	0.80	0.40	0	1	668
Personality	0.84	0.37	0	1	668
Experience	0.23	0.42	0	1	668
Knowledge	0.52	0.50	0	1	668
Social skills	0.58	0.50	0	1	668
Administrative skills	0.52	0.50	0	1	668
Leadership competencies	0.64	0.48	0	1	668
Interests	0.20	0.40	0	1	668
Work values	0.48	0.50	0	1	668
Motivation	0.57	0.50	0	1	668
Number of characteristics assessed with tests	5.36	2.40	0	10	668
Data & Testing Security[a]					
Restrict access to testing data	0.98	0.13	0	1	607
Disaster recovery plans in place	0.66	0.47	0	1	337
Have physical security	0.86	0.35	0	1	473
Encrypt testing data	0.78	0.41	0	1	422
Regular testing data backups	0.87	0.34	0	1	460
Protect testing data with firewall and password	0.95	0.23	0	1	539
Use at least one data protection method	0.96	0.18	0	1	652
Use at least one security measure for paper & pencil tests	0.88	0..33	0	1	509
Use at least one security measure for unsupervised computerized tests	0.55	0.50	0	1	507
Use at least one security measure for supervised computerized tests	0.52	0.50	0	1	620
Frequency of disqualifying applicants for cheating[d]					
Paper & pencil tests	1.73	0.91	1	5	477
Supervised computerized tests	1.60	0.86	1	5	338
Unsupervised computerized tests	1.48	0.77	1	5	462

[a]Responses dummy coded as 0 = no, 1 = yes
[b]Response scale is 1 = never, 2 = rarely (1-20% of the time), 3 = occasionally (21-50% of the time), 4 = frequently (51-80% of the time), 5 = Almost always or always (81-100% of the time)
[c]Responses dummy coded as 1 = process is same for all countries, 2 = different processes used across countries, 3 = mixture of custom and standardized processes across countries
[d]Response scale is 1 = never, 2 = very rarely, 3 = somewhat rarely, 4 = somewhat frequently, 5 = very frequently

hiring procedure (M = 5.36, SD = 2.40), suggesting that many companies appear to be evaluating multiple constructs as part of their applicant selection process. With respect to security measures related to hiring assessments, the use of procedures designed to protect data (i.e., measures taken

to ensure confidentiality, prevent loss of/tampering with data, etc.) were far more common (96.5% of organizations) than the use of procedures designed to minimize cheating/dishonest responding by test takers (i.e., standardized testing protocols, controlling applicant access to testing materials, etc.; 88% for paper-and-pencil testing, 52% for supervised computerized testing). The percentage of organizations which employ proctoring as a means of enhancing test taking security was even smaller. Only 42.4% (n = 283) and 24.1% (n = 161) of organizations reported that all of their paper-and-pencil or computerized tests, respectively, were administered in supervised settings. Supervision of all tests was significantly more common with the use of paper-and-pencil than computerized assessments ($t(667) = 8.76$, $p < .001$).

Variance Components and Intraclass Correlation

In the present study, all outcome variables were measured at the organizational level. The outcome variables may thus exhibit variability due to two sources: differences between organizations within countries and differences between countries. Multilevel models can be used to partition variance in each outcome measure according to these different sources as well as attempt to account for variability at both levels of analysis. For the present analyses, interest is in whether the four societal culture predictor variables measured at the country-level account for variability in organizational-level outcome variables. Consequently, it is prudent to first examine whether sufficient between-country variability exists in these outcome variables.

We fit an initial regression model including no country-level predictors to estimate the degree of variability at the organization versus country levels of analysis, as well as the intraclass correlation $ICC(1)$ for each outcome variable. The index $ICC(1)$ provides an estimate of the proportion of variance in an outcome that can be explained by group membership. Table 4 summarizes the variance components, ICCs, and sample sizes at the organizational and country

Table 4. Intraclass Correlation (ICC) Estimates for Outcome Variables

	Country VC	Error VC	ICC(1)	ICC(2)	Number Obs	Number Countries
Use of tests for selecting entry-level management employees	0.29	3.29	0.08	0.68	1064	24
Use of customized testing procedures	0.47	3.29	0.12	0.77	668	23
Use of unsupervised paper-and-pencil tests	0.43	3.29	0.12	0.75	668	23
Use of unsupervised computerized tests	0.38	3.29	0.10	0.72	668	23
Frequency of feedback provided to applicants	0.54	1.37	0.28	0.90	657	23
Standardization of multinational testing practices	0.00	0.57	0.00	0.01	202	21
Assessed with tests: Ability	0.01	3.29	0.00	0.05	668	23
Assessed with tests: Personality	0.61	3.29	0.16	0.81	668	23
Assessed with tests: Experience	0.20	3.29	0.06	0.59	668	23
Assessed with tests: Knowledge	0.45	3.29	0.12	0.76	668	23
Assessed with tests: Social Skills	0.05	3.29	0.01	0.26	668	23
Assessed with tests: Administrative Skills	0.31	3.29	0.09	0.69	668	23
Assessed with tests: Leadership Competencies	0.47	3.29	0.12	0.77	668	23
Assessed with tests: Interests	0.31	3.29	0.09	0.69	668	23
Assessed with tests: Work Values	0.22	3.29	0.06	0.60	668	23
Assessed with tests: Motivation	0.23	3.29	0.07	0.62	668	23
Sum of different characteristics assessed with tests	0.33	5.47	0.06	0.58	668	23
Data protection: Restricted access	0.00	3.29	0.00	0.00	607	23
Data protection: Disaster recovery plans	0.16	3.29	0.05	0.49	337	20
Data protection: Physical security	0.54	3.29	0.14	0.78	473	22
Data protection: Encryption	0.24	3.29	0.07	0.61	422	21
Data protection: Data backups	0.69	3.29	0.17	0.82	460	21
Data protection: Firewall and password protection	0.41	3.29	0.11	0.74	539	23
Data protection: Indicated use of any data protection method	0.21	3.29	0.06	0.59	652	23
Use of at least one paper and pencil security measure	0.40	3.29	0.11	0.73	509	23
Use of at least one unsupervised computerized measure	0.05	3.29	0.02	0.27	507	23
Use of at least one supervised computerized measure	0.13	3.29	0.04	0.48	620	23

Frequency of disqualifying applicants for cheating	0.07	0.77	0.08	0.67	477	23
Frequency of disqualifying applicants for cheating on supervised computerized tests	0.07	0.69	0.09	0.69	338	21
Frequency of disqualifying applicants for cheating on unsupervised computerized tests	0.21	0.50	0.30	0.91	462	23

Note. VC = variance component. For binary outcomes, intraclass correlation computed per Snijders and Bosker (1999; p. 224).

levels for all outcome variables. Overall, the results reveal that the *ICC*(1) values were non-zero, but generally small. This indicates that there was more variance in the measured outcomes attributable to differences between organizations rather than differences between countries. However, Cohen, Cohen, West, and Aiken (2003) note that even an *ICC*(1) as low as .05 can have a significant effect on the results of statistical analyses that do not control for clustering, and thus the use of multilevel modeling in this instance is warranted.

Performance Orientation Hypotheses

Hypothesis 1 predicted that performance orientation would be positively related to (1a) adopting the use of tests for hiring purposes, (1b) adopting procedures that elicit behaviors and skills that are clearly relevant to job performance (e.g., knowledge, skills), and (1c) providing feedback to applicants.

Table 5 contains results pertaining to Hypotheses 1a-1c. Hypothesis 1a was not supported; performance orientation was not related to the use of tests (b = -0.13, SE = 0.39, p = 0.74). Hypothesis 1b addressed the relationship between performance orientation and endorsing the use of tests that assess abilities, knowledge, social skills, and administrative skills. Performance orientation was significantly related to use of tests that assess knowledge (b = 1.06, SE = 0.43, p = 0.01), such that organizations in highly performance-oriented countries (e.g., New Zealand, Singapore, Hong Kong) were more likely to indicate testing for knowledge than were

Table 5. Mixed-effects Model Estimates for the Regression of Study Outcomes on Country-level Cultural Practices

	Intercept		Future Orientation		Performance Orientation		Uncertainty Avoidance		Cultural Tightness		Number Countries	Number Obs
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*		
Use tests for selecting entry-level management employees[1]	***0.53***	0.13	-0.22	0.47	-0.13	0.39	***0.82***	0.33	-0.07	0.07	23	1062
Use customized testing procedures	***0.33***	0.14	-0.75	0.58	***0.79***	0.42	-0.14	0.35	0.16	0.09	22	666
Use unsupervised paper & pencil tests[1]	***-0.80***	0.13	***-2.07***	0.57	0.10	0.37	***0.88***	0.35	0.03	0.09	22	666
Use unsupervised computerized tests[1]	***0.56***	0.16	0.03	0.58	-0.58	0.47	0.63	0.41	-0.11	0.09	22	666
Frequency of feedback to applicants	***3.85***	0.15	0.48	0.51	***-1.14***	0.43	0.58	0.37	***-0.17***	0.08	22	655
Standardization of multinational testing practices	***2.18***	0.06	0.02	0.23	-0.11	0.15	0.04	0.16	0.01	0.04	20	201
Assessed with tests: Ability	***1.37***	0.10	-0.23	0.40	0.22	0.28	-0.04	0.27	-0.07	0.08	22	666
Assessed with tests: Personality	***1.83***	0.22	0.22	0.77	-0.71	0.62	0.51	0.53	-0.17	0.11	22	666
Assessed with tests: Experience	***-1.33***	0.15	-0.29	0.53	0.70	0.42	-0.52	0.37	0.02	0.09	22	666
Assessed with tests: Knowledge[1]	-0.14	0.16	-0.32	0.53	***1.06***	0.43	***-0.79***	0.36	***0.24***	0.09	22	666
Assessed with tests: Social Skills	***0.32***	0.08	0.02	0.33	***-0.68***	0.24	0.28	0.23	-0.05	0.07	22	666
Assessed with tests: Administrative Skills	-0.20	0.18	-0.13	0.65	-0.70	0.51	0.08	0.44	-0.07	0.10	22	666
Assessed with tests: Leadership Competencies	***0.49***	0.16	0.59	0.57	-0.66	0.46	0.47	0.41	-0.07	0.09	22	666
Assessed with tests: Interests	***-1.55***	0.18	-0.83	0.62	0.72	0.51	-0.14	0.47	0.01	0.10	22	666
Assessed with tests: Work Values	-0.13	0.14	-0.94	0.53	-0.26	0.40	0.67	0.36	-0.07	0.08	22	666
Assessed with tests: Motivation	***0.28***	0.10	-0.44	0.43	-0.49	0.33	***0.99***	0.31	-0.03	0.07	22	666
Number of characteristics assessed with tests	***5.07***	0.22	-0.17	0.76	-0.55	0.61	0.12	0.54	-0.04	0.12	22	666
Data protection: Restricted access	***4.16***	0.34	-1.21	1.35	-0.16	1.01	0.97	0.88	-0.15	0.19	22	605
Data protection: Disaster recovery	***0.61***	0.14	0.91	0.62	-0.44	0.38	-0.16	0.38	-0.01	0.10	19	336

plans												
Data protection: Physical security[1]	***1.89***	0.19	***1.64***	0.72	***-1.25***	0.60	0.19	0.53	0.09	0.14	21	472
Data protection: Encryption	***1.21***	0.19	0.58	0.68	-0.58	0.56	0.15	0.48	-0.04	0.12	20	421
Data protection: Data backups[1]	***2.02***	0.22	***1.63***	0.79	0.65	0.56	-0.57	0.50	-0.20	0.13	20	459
Data protection: Firewall and password protection	***3.14***	0.25	***1.98***	0.83	-0.18	0.65	-0.09	0.53	-0.30	0.15	22	538
Data protection: Use at least one data protection method	***3.50***	0.25	0.94	0.81	-0.41	0.66	-0.50	0.52	-0.28	0.15	22	650
Use at least one security measure for paper & pencil tests	***2.10***	0.15	-1.03	0.71	***1.26***	0.39	-0.39	0.50	0.25	0.14	22	507
Use at least one security measure for unsupervised computerized tests	***0.19***	0.09	-0.73	0.42	***0.58***	0.26	0.37	0.30	***-0.19***	0.09	22	506
Use at least one security measure for supervised computerized tests[1]	0.09	0.08	***-0.72***	0.35	0.43	0.24	-0.22	0.25	0.14	0.07	22	618
Frequency of disqualifying applicants for cheating on paper & pencil tests	***1.75***	0.09	-0.40	0.30	0.27	0.24	-0.05	0.21	0.00	0.05	22	475
Frequency of disqualifying applicants for cheating on supervised computerized tests	***1.62***	0.08	-0.29	0.31	0.17	0.24	-0.27	0.21	0.03	0.05	20	336
Frequency of disqualifying applicants for cheating on unsupervised computerized tests	***1.59***	0.13	0.07	0.41	0.33	0.36	-0.38	0.32	0.06	0.07	22	461

Note. Est = Estimate; SE = Standard Error. Estimates in bold italicized font are significant at the $p < 0.05$ level.

organizations in low performance orientation countries (e.g., Greece, Russia, Portugal). However, performance orientation was negatively related to use of tests that assess social skills (b = -0.68, SE = 0.24, $p < 0.01$) and not related to the use of tests for assessing ability (b = 0.22, SE = 0.28, p = 0.43) or administrative skills (b = -0.70, SE = 0.51, p = 0.17). Thus, results were generally not supportive of Hypothesis 1b. Hypothesis 1c was also not supported. Contrary to what was hypothesized, a significant negative relationship was found between performance orientation and frequency with which feedback is provided to applicants (b = -1.14, SE = 0.43, p = 0.02). Thus, organizations in highly performance oriented countries were less likely to indicate that they provide feedback to applicants.

Future Orientation Hypotheses

Hypotheses 2a-2c addressed outcomes related to future- orientation. Namely, it was hypothesized that organizations in highly future-oriented countries (e.g., Singapore, the Netherlands, Denmark) would be more likely to (2a) adopt the use of tests for hiring purposes, (2b) test for a greater variety of different attributes within the organization's hiring process, and (2c) adopt customized or self-developed tools as opposed to off-the-shelf solutions than would organizations in low future-orientated cultures (e.g., Russia, Italy, New Zealand).

Table 5 summarizes results pertaining to Hypotheses 2a-2c, none of which were supported. Future orientation was not related to the use of tests (b = -0.22, SE = 0.47, p = 0.64), the number of different types of procedures that organizations endorsed using (b = -0.17, SE = 0.76, p = 0.82), and use of customized testing procedures (b = -0.75, SE = 0.58, p = 0.20).

Uncertainty Avoidance Hypotheses

Hypotheses 3a-3d addressed outcomes specific to uncertainty avoidance. Namely, it was hypothesized that organizations in high uncertainty avoidant countries (e.g., Sweden, Germany,

Denmark) would be more likely to (3a) test for a greater variety of different attributes within the organization's hiring process, (3b) use more structured testing procedures (e.g., greater standardization of procedures on a global basis), (3c) use a larger number of test security measures, and (3d) use unproctored modes of test administration than would organizations in low uncertainty avoidant countries (e.g., Russia, Greece, Turkey).

Table 5 contains results pertaining to Hypotheses 3a-3d. No evidence was found that uncertainty avoidance was related to the number of different attributes assessed ($b = 0.12$, $SE = 0.54$, $p = 0.82$) or the amount of standardization employed in testing practices globally ($b = 0.04$, $SE = 0.16$, $p = 0.82$). With regard to test security methods (Hypothesis 3c), uncertainty avoidance was not related to adoption of any of several data protection methods (e.g., use of physical security, encryption methods, firewalls; $b = -0.50$, $SE = 0.52$, $p = 0.34$). Finally, uncertainty avoidance was related to organizations' inclination to use unproctored paper-and-pencil tests ($b = 0.88$, $SE = 0.35$, $p = 0.01$), but not unproctored computerized tests ($b = 0.63$, $SE = 0.41$, $p = 0.12$). Thus, whereas no evidence was found for Hypotheses 3a-3c, mixed support was found for Hypothesis 3d.

Tightness-looseness Hypotheses

Hypotheses 4a-4e address outcomes specific to tightness norms. Namely, it was hypothesized that organizations in culturally tight countries would (4a) be more likely to adopt the use of tests for hiring purposes, (4b) be more likely to test for attributes associated with personality, work styles, and other personal characteristics, (4c) be less likely to test for attributes associated with knowledge, skills, and abilities, (4d) be more likely to use test security methods, and (4e) be less likely to conduct unproctored testing than would organizations in culturally loose countries.

Table 5 contains results pertaining to Hypotheses 4a-4e. Results did not provide support for Hypothesis 4a. Organizations in culturally tight countries (e.g., India, Singapore, Indonesia) were not more likely to use testing practices than were organizations in culturally loose countries (e.g., the Netherlands, Brazil, New Zealand; b = -0.07, SE = 0.07, p = 0.34). Hypotheses 4b and 4c were also generally not supported. Cultural tightness was positively related to the use of tests to assess knowledge (b = 0.24, SE = 0.09, p = 0.01), but none of the remaining effects for Hypothesis 4c were significant. Specifically, cultural tightness was not related to the use of tests to assess ability (b = -0.07, SE = 0.08, p = 0.35), personality (b = -0.17, SE = 0.11 , p = 0.12), experience (b = 0.02, SE = 0.09 , p = 0.82), social skills (b = -0.05, SE = 0.07 , p = 0.40), administrative skills (b = -0.07, SE = 0.10 , p = 0.50), leadership (b = -0.07, SE = 0.09 , p = 0.46), interests (b = 0.01, SE = 0.10 , p = 0.94), work values (b = -0.07, SE = 0.08 , p = 0.43), or motivation (b = -0.03, SE = 0.07 , p = 0.66).

Results were generally not supportive of Hypothesis 4d. There was no evidence that cultural tightness was related to the use of test security measures associated with data protection (b = -0.28, SE = 0.15, p = 0.07). Evidence was not found for a relationship between cultural tightness and use of security measures for paper-and-pencil tests (b = 0.25, SE = 0.14, p = 0.07). The relationship with security measures for unproctored computerized tests was negative (b = -0.19, SE = 0.09, p = 0.04), although no relationship was found involving security measures for proctored computerized tests (b = 0.14, SE = 0.07, p = 0.05). No relationship was found involving disqualification of applicants for cheating on paper-and-pencil tests (b = 0.00, SE = 0.05, p = 0.96), proctored computerized tests (b = 0.03, SE = 0.05, p = 0.56), or unproctored computerized tests (b = 0.06, SE = 0.07, p = 0.40).

Finally, results were not supportive of Hypothesis 4e. Reported use of unproctored paper-and-pencil tests was not related to cultural tightness for either paper-and-pencil tests ($b = 0.03$, $SE = 0.09$, $p = 0.76$) or unproctored computerized tests ($b = -0.11$, $SE = 0.09$, $p = 0.24$).

Additional Considerations

Though not hypothesized, Table 5 shows several significant relationships related to the use of data protection and security measures in testing. However, there was no clear pattern to these results. Consequently, we conducted additional analyses to see whether any organization-level variables were relevant to outcome prediction. The relationships between two additional organizational-level factors—sector (public/private) and organizational size—with reported selection assessment practices were examined by regressing the outcomes on the organization-level predictors. To accommodate clustering by country, we allowed the intercepts to vary in the models. In relation to sector, private sector organizations did not significantly differ from public sector organizations on their reported use of tests during selection ($b = -0.33$, $SE = 0.17$, $p = 0.05$), the use of customized vs. off-the-shelf assessments ($b = -0.36$, $SE = 0.21$, $p = 0.08$), or the provision of feedback following selection assessments ($b = -0.05$, $SE = 0.11$, $p = 0.63$).

Differences between public and private sector organizations were found with respect to the format of tests used, the use of unproctored testing, and the use of test-taking security measures. Private sector organizations were less likely to report administering paper-and-pencil tests than were public sector organizations ($b = -0.65$, $SE = 0.25$, $p < 0.01$), but were more likely to administer computerized tests ($b = 0.58$, $SE = 0.29$, $p = 0.04$). Concerning unproctored testing, private sector organizations were significantly more likely than those in the public sector to administer unproctored paper-and-pencil tests, ($b = 0.46$, $SE = 0.23$, $p = 0.04$) and administer computerized unproctored tests ($b = 0.92$, $SE = 0.21$, $p < 0.01$). With respect to test security

practices, there was some indication that public sector organizations reported use of stricter security measures than private sector companies. In particular, public sector organizations were more likely to report adopting a test-taking security precaution than were private sector companies when using paper-and-pencil tests *(b* = 0.89, *SE* = 0.42, *p* = 0.04) and proctored computerized tests (*b* = 0.69, *SE* = 0.21, *p* < 0.01), whereas no difference was found between public and private sector organizations with regarding to reported use of security precautions for unproctored computerized tests (*b* = 0.39, *SE* = 0.24, *p* = 0.10).

The reported number of employees in an organization exhibited only one significant relation with the assessment practices of interest. A trend was observed in the propensity to use tests such that larger organizations reported a higher likelihood of using tests than smaller organizations (*b* = .22, *SE* = 0.10, *p* = 0.02).

Finally, a reviewer noted that a respondent's country location may differ from the multinational headquarters (HQ) location, and thus the societal culture most influencing the organization might be more appropriately represented by the country of the HQ rather than that of the survey respondent. To address this concern, we reanalyzed the data with HQ country substituted for country where applicable. The reanalysis led to no changes in lack of support for hypotheses, except that the previously stated mixed support for H3d (uncertainty avoidance significantly related to use of unsupervised paper-and-pencil tests) was now totally non-supportive. Thus, considering the HQ country for mulitnationals rather than the country of the respondent did not in any way change our conclusions.

DISCUSSION

Many authors have speculated how societal culture might influence selection practices (see Steiner, 2012 and Caligiuri & Paul, 2010 for reviews). However, using a large sample of HR

respondents, we found little evidence that societal cultural characteristics are related to testing practices. These results are consistent with Ryan et al.'s (1999) study that also found very few connections between Hofstede's cultural values and testing practices. Coupled with the number of studies that have established cross-cultural equivalence of measures (e.g., Bartram, 2013) and those showing cross-cultural similarity in applicant reactions (see Steiner, 2012 for review), the empirical research to date suggests that selection practices are likely to generalize across cultures.

Why might there be little connection between societal cultural characteristics and adopted selection practices? First, scholars have repeatedly pointed out the flawed thinking behind assuming societal cultural differences translate into differences in individual or even group behavior (e.g., Brewer & Venaik, 2012; Gelfand, Erez, & Aycan, 2007). For example, McSweeney (2013) reminds us of the ecological fallacy of inferring that characteristics of an aggregate (society) also describe entities at lower levels (organizations in those societies). Researchers in organizational behavior have emphasized this fallacy with regard to assuming individual endorsement of societal values (e.g., not all individuals in China, a collectivist society, endorse collectivist values; see Gerhart & Fang, 2005 and Oysernam, Coon, & Kemmelmeier, 2002 for reviews showing that national culture values explain only 1-4% of individual cultural value endorsement). However, there is still a tendency to believe that organizational level practices, like selection tool use, will reflect societal norms in some fashion. McSweeney notes that this tendency to view the macro as creating the micro—in terms of societal values causing behavior at a sub-national level—is not an uncommon belief; Gelfand and colleagues (2007) have repeatedly called for systematic examinations of cross-level relationships to avoid this levels of analysis confusion. We hope that this study's "non-findings" with regard to societal

cultural characteristics affecting organizational practices represents a step forward in thinking about selection practices globally, and suggest similar possibilities for other HR practices.

Second, while cultural norms represent a top-down, "constraining" influence on what should be appropriate in a given context, the range of appropriate behaviors established by societal culture may be too large to exhibit impact on processes as "mundane" as selection testing. That is, the typical procedures relevant in selection contexts may not be such that they clash with any of the limits implied by certain societal cultural norms. How societal cultural values and practices relate to compensation, work-family policies, performance evaluation systems, and other HR systems may also warrant systematic exploration.

Third, one important future direction may be examination of the extent to which the fit between *organizational* cultural values and practices and *societal* cultural values and practices influences selection practices. That is, an organization's culture is a strong influence on its HR practices, but perhaps it is only when organizational culture and societal culture clash that societal cultural norms become very salient in selection system design.

Fourth, it is important to remember that we did find considerable variability within country in practices (i.e., our lower ICC values), suggesting that this area of HR practice is one that differentiates organizations. Theoretical development and research on the sources of variability can help advance thinking about selection but also about HR practice diffusion more broadly. For example, what leads an organization to adopt a technological advance in assessment (e.g., use of computer adaptive testing, gamification) versus not given similarities in applicant pools and resources? Czarniawska and colleagues (Czarniawska & Joerges, 1995, 1996; Czarniawska & Sevon, 2005) note that ideas and innovations are not invariant when "traveling" globally; they are translated or adapted by each user (e.g., the idea of a "family- friendly"

organization may not mean the same thing in different locales, even though all are embracing the idea of a change in policies, Frenkel, 2005). Ideas that travel widely around the globe are those that connect to a similar desire or need on the part of organizations. As an example, innovations in assessment that appear to be travelling globally may all connect with organizational needs for more efficient hiring processes; these are not cases of the same exact tool adopted in the same exact way, but "translated" versions of ideas to increase efficiency. Qualitative approaches that follow the translation of an assessment innovation globally would yield useful insights into how and why certain selection ideas get wider, global traction (Czarniawska & Sevon, 2005).

More specific investigation of globalization forces might also be useful; that is, Meyer (2002) suggested that there are multiple pathways through which globalization changes organizations. In the case of cultural practices, if globalization weakens the control and legitimacy of the national community (Meyer, 2002), the influence of societal cultural practices on HR practices may lessen. However, globalization also expands markets, so an HR innovation with a greater market to tap may contribute to the spread of ideas.

Note that the above points are useful directions for all culture-comparative studies, not just for those specifically on selection method diffusion. That is, all culture-comparative studies should avoid ecological fallacies, consider culture as a top-down influence on behavior, examine the fit between organizational and societal culture, and adopt frameworks to understand how ideas and innovations travel. Such advice will advance our general understanding of when and why societal cultural norms might affect organizational practices.

Practical Implications

The conclusions practitioners can draw regarding societal culture's lack of influence on selection practices are tempered by several very important caveats. First, while societal cultural

characteristics may not influence the adoption of certain testing practices, this does not mean any particular assessment tool is automatically culturally transportable. It is still important to ensure that a selection test is psychometrically equivalent across settings. Second, acceptability of an assessment tool based on societal cultural norms (or even psychometric equivalence) does not necessarily mean that its implementation in a particular context is a good business decision. Ryan and Tippins (2009) provide an extensive discussion of the many practical hurdles (e.g., lack of available technology, lack of available administrative personnel, legal differences, union/work council objections) that may suggest the need to go slowly in importing tests from one context to another or to develop "work arounds" or variations in specific countries (see Ryan et al., 2003 for further examples). Third, testing practices did vary across organizations, and such variability in practice may correlate with national context, due to legal and economic variability. For example, data protection will be higher when laws require it and the use of computerized assessments and sophisticated advances in online testing (e.g., adaptive testing) will be more prominent in locations where such administration is economically feasible and supportable by technological infrastructure.

Strengths and Limitations

This study has a number of strengths: a large sample of HR professionals from 23 countries responded to the survey, the sampling design sought to maximize variability in key cultural characteristics reported in GLOBE (House et al., 2004), and we invested in quality translation processes to ensure that the survey items were well understood. However, there are several limitations stemming from the challenges of undertaking such a large-scale effort.

First, as in any study, decisions were made regarding item focus so as to keep the survey at a reasonable length. While objective practices can be assessed with single item measures (e.g.,

"Do you use tests or not?"), we certainly may have missed nuances of practice. Second, we made decisions in our sampling frame as to which societal cultural practices to focus on based on our review of the literature and theoretical rationale for hypotheses. While we feel our exclusion of other cultural characteristics was justified based on a lack of clear connections to selection practices (e.g., prior research does not support individualism/collectivism as connected to practice use), other cultural frameworks may cast a different light on selection practices. Note that the cultural practices within the GLOBE framework are not orthogonal. Cultural practice scores on the performance orientation, future orientation, and uncertainty avoidance dimensions are significantly ($p < .05$) and positively inter-related (performance orientation-future orientation $r = .63$, performance orientation-uncertainty avoidance $r = .58$, and future orientation-uncertainty avoidance $r = .76$; see Hanges, 2004b, Table A.1, p. 734). Third, we were limited to a single respondent for each organization, which has limitations in measuring HR practices (Wright et al., 2001).

A fourth potential limitation concerns a lack of power and small between-country variability (i.e., low *ICC*(1) values) that may inhibit the ability to detect small effects. With respect to the former, power is likely to be limited in any study examining relationships between testing practices and country-level characteristics because of the practical challenges inherent in data collection in developing countries (e.g., many studies similar to ours may only have data from three to ten countries), the inherent ceiling in sampling countries as a unit of analysis (i.e., there are only a limited number of countries in the world), and the distal nature of the likely relationship between national culture and organizational testing policies. Note that our focus was on culture rather than country so as to avoid limitations of small Ns per country; however, greater power might have led to possibly detecting small effects. Given the complexity of our

analyses, we used Monte Carlo simulation as opposed to analytic (formula-based) procedures to obtain estimates for power. [5]We estimated power for both types of models that were fit in our study (linear and logistic mixed-effects) by choosing one outcome of each type modeled in our study and using it as the basis for generating simulated data. *Power* was calculated as the proportion of replications where the observed *p*-value was less than the nominal alpha level of 0.05 (e.g., Feiveson, 2002; Gelman & Hill, 2007). For instance, across the 5,000 replicated simulations for the linear model with feedback frequency as an outcome, the proportion of instances where Future Orientation was significant was 0.16. Aside from the intercept term, the highest observed power for feedback frequency was for Performance Orientation (0.67) and Tightness (0.61). Power for Uncertainty Avoidance (0.33) and Future Orientation (0.16) was, comparably speaking, much lower. For the logistic model with the outcome of assessing experience, power associated with the slopes for each of the predictors was 0.15 for Future Orientation, 0.33 for Performance Orientation, 0.46 for Uncertainty Avoidance, and 0.22 for Tightness.

With respect to the degree of between-country variability observed in our sample, the presence of low *ICC*(1) values made it unlikely that we would find support for our hypotheses. However, we do not see this as a limitation, but rather a substantive finding of interest. Given our strategy of sampling level-2 units (i.e., countries) based on their cultural practice scores, the low *ICC*(1) values are strongly consistent with the conclusion that there is little variance in selection practices attributable to differences in the endorsement of societal culture characteristics. A fifth limitation is that running a large number of statistical tests raises the probability of making a Type I error. If we were to correct the .05 alpha level used (e.g., adopt $p < .002$ with a Bonferroni correction), the number of non-significant results would further increase (see Table 5). However,

[5] For a detailed description of the power analysis simulation, contact the authors.

this reinforces the main conclusion of this study—that societal cultural values and practices appear mostly unrelated to organizational testing practices.

Finally, we limited our focus to testing practices rather than all aspects of selection practices (e.g., recruitment, interviewing) as this is the area we felt had the most dramatic advancements and changes since the Ryan et al. (1999) study; the influence of cultural characteristics on other aspects of hiring processes is worthy of further study.

Conclusion

In this study, we found little evidence that societal cultural characteristics are associated with testing practices in organizations. While lack of support for our hypotheses could have been due to limitations in methodology (e.g., sampling plan, choice of items), the cumulative body of evidence regarding a lack of strong influence of societal cultural characteristics on selection practice acceptability (Ryan et al., 1999; Ryan et al., 2009; current study), coupled with the knowledge that such cross-level hypotheses appear to draw most of their conceptual support from ecological fallacies regarding causal connections across levels of analysis, lead us to conclude that variability in selection practices is likely not strongly associated with societal culture.

Acknowledgement

We wish to acknowledge the assistance of Casey Wood in retrieving information for this study.

References

Anderson, N., Salgado, J. F., & Hulsheger, U. R. (2010). Applicant reactions in selection: Comprehensive meta-analysis into reaction generalization versus situational specificity. *International Journal of Selection and Assessment, 18*(3), 291-304.

Ashkanasy, N., Gupta, V., Mayfield, M.S., & Trevor-Roberts, E. (2004). Future orientation. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta, *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 282-342). Thousand Oaks, CA: Sage.

Atwater, L., Wang, M., Smither, J. W., & Fleenor, J. W. (2009). Are cultural characteristics associated with the relationship between self and others' ratings of leadership? *Journal of Applied Psychology, 94*(4), 876-886.

Bartram, D. (2001). International guidelines for test use. *International Journal of Testing, 1*(2), 93–114

Bartram, D. (2008). Global norms: Towards some guidelines for aggregating personality norms across countries. *International Journal of Testing, 8*(4), 315-333.

Bartram, D. (2013). Scalar equivalence of OPQ32: Big five profiles of 31 countries. *Journal of Cross-Cultural Psychology, 44*(1), 61-83.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0-4. http://cran.r-project.org/package=lme4.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349-381). San Francisco, CA: Jossey-Bass.

Brewer, P., & Venaik, S. (2012). On the misuse of national culture dimensions. *International Marketing Review*, *29*(6), 673-683.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross cultural research* (pp. 137-164). Thousand Oaks, CA: Sage.

Caligiuri, P., & Paul, K. (2010). Selection in multinational organizations. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of employee selection* (pp. 781-800). New York: Routledge.

Clark, T. (1993). Selection methods used by executive search consultancies in four European countries: A survey and critique. *International Journal of Selection and Assessment, 1*(1), 41-49.

Cohen J., Cohen P., West S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed). Mahwah, NJ: Lawrence Erlbaum Associates.

Czarniawska, B. & Joerges, B. (1995). Winds of organizational change: how ideas translate into objects and actions. *Research in the Sociology of Organizations, 13*, 171-209.

Czarniawska, B. & Joerges, B. (1996). Travel of ideas. In B. Czarniawska & G. Sevin (Eds). *Translating organizational change*. Boston: DeGruyer. 13-48.

Czarniawska, B. & Sevon, G. (Eds.). (2005). *Global ideas: How ideas, objects, and practices travel in the global economy.* Denmark: Liber and Copenhagen Business School Press.

Dai, G., Han, K., Hu, H., & Colarelli, S. M. (2010). Cultural differences and measurement invariance of selection tools: A case of examining Chinese NEO PI-R conscientiousness scale. *Journal of Chinese Human Resource Management, 1*(2), 95-114

Farr, J.L., & Tippins, N.T. (Eds.), *Handbook of employee selection*. New York: Routledge.

Fell, C.B. & Konig, C.J. (2016). Cross-cultural differences in applicant faking on personality tests: a 43-naiton study. *Applied Psychology: An International Review, 65*, 671-717.

Fell, C. B., König, C. J., & Kammerhoff, J. (2015). Cross-cultural differences in the attitude toward applicants' faking in job interviews. *Journal of Business and Psychology,* doi:http://dx.doi.org/10.1007/s10869-015-9407-8

Frenkel, M. (2005). Something new, something old, something borrowed: the cross-national translation of the "family friendly organization" in Israel. In Czarniawska, B. & Sevon, G. (Eds.). (2005). *Global ideas: How ideas, objects, and practices travel in the global economy*. Denmark: Liber and Copenhagen Business School Press. 147-166.

Gefland, M.J., Erez, M., & Aycan, Z. (2007). Cross-cultural organizational behavior. *Annual Review of Psychology*, *58*, 479-514.

Gelfand, M. J., Leslie, L. M., & Fehr, R. (2008). To prosper, organizational psychology should...adopt a global perspective. *Journal of Organizational Behavior, 29*(4), 493-517.

Gelfand, M. J., Nishii, L. H., & Raver, J. L. (2006). On the nature and importance of cultural tightness-looseness. *Journal of Applied Psychology, 91*(6), 1225-1244.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., … Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science, 332*, 1100-1104.

Gerhart, B., & Fang, M. (2005). National culture and human resource management: Assumptions and evidence. *The International Journal of Human Resource Management*, *16*(6), 971-986.

Hambleton, R. K. (2005). Issues, designs and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Speilberger (Eds.), *Adapting educational and psychological tests of cross-cultural assessment* (pp. 3-38). Mahwah NJ: Erlbaum.

Hanges, P.J. (2004a). Response bias correction procedure used in GLOBE (Appendix B). In R.J. House, P.J. Hanges, M. Javidan, P.W. Dorfman, & V. Gupta, *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 737-751). Thousand Oaks, CA: Sage.

Hanges, P. J. (2004b). Societal-level correlations among GLOBE societal culture scales (Appendix A). In R.J. House, P.J. Hanges, M. Javidan, P.W. Dorfman, & V. Gupta, *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 733-736). Thousand Oaks, CA: Sage.

Hofstede, G. (1980). Culture's consequences: International differences in work-related values. Beverly Hills, CA: Sage.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P.W., & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.

Huo, Y. P., Huang, H. J., & Napier, N. K. (2002). Divergence or convergence: A cross-national comparison of personnel selection practices. *Human Resource Management, 41,* 31-44.

International Test Commission (2010). *International test commission guidelines for translating and adapting tests.* http://www.intestcom.org

Javidan, M. (2004). Performance orientation. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 239-281). Thousand Oaks, CA: Sage.

Javidan, M. & House, R.J. (2001). Cultural acumen for the global manager: Lessons from Project GLOBE. *Organizational Dynamics, 29*, 289-305.

Javidan, M., House, R.J., & Dorfman, P. W. (2004). A non technical summary of GLOBE findings. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 29-48). Thousand Oaks, CA: Sage.

Krause, D. E., & Thornton, G. C. (2009). A cross-cultural look at assessment center practices: Survey results from western Europe and North America. *Applied Psychology: An International Review, 58*, 557-585.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11,* 815-852.

Lim, C., Chavan, M., & Chan, C. (2014). "Culture"—The elephant in the room in structured behavioral selection interview.*International Journal of Intercultural Relations, 42*, 1-24. doi:http://dx.doi.org/10.1016/j.ijintrel.2014.07.004

Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods, 7,* 361-388.

McSweeney, B. (2013). Fashion founded on a flaw: The ecological mono-deterministic fallacy of Hofstede, GLOBE, and followers. *International Marketing Review, 30*(5), 483 – 504.

Meyer, J. (2002). Globalization and the expansion and standardization of management. In K. Sahlin-Andersson & L. Engwall (Eds.). *The expansion of management knowledge: carriers, flows, and sources.* CA: Stanford Business Books. 33-46.

Ott-Holland, C. J., Huang, J. L., Ryan, A.M., Elizondo, F., & Wadlington, P. L. (2013). Culture and vocational interests: The moderating role of collectivism and gender egalitarianism. *Journal of Counseling Psychology, 60*(4), 569-581.

Oysernam, D., Coon, H.M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses, *Psychological Bulletin*, *128*(1), 3-72.

R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.

Ryan, A.M., Inceoglu, I., Bartram, D., Golubovich, J., Grand, J., Reeder, M., Derous, E., Nikolaou, I., & Yao, X. (2015). Trends in Testing: Highlights of a Global Survey. To appear in I. Nikolaou & J. K. Oostrom (Eds.). *Employee recruitment, selection and assessment: Contemporary issues for theory and practice.* Psychology Press. Taylor & Francis.

Ryan, A.M., Boyce, A. S., Ghumman, S., Jundt, D., Schmidt, G., & Gibby, R. (2009). Going global: Cultural values and perceptions of selection procedures. *Applied Psychology: An International Review, 58,* 520-566.

Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology, 52*(2), 359-391.

Ryan, A. M., & Tippins, N. (2009). *Designing and implementing global selection systems.* Malden, MA: Wiley Blackwell.

Ryan, A. M., Wiechmann, D., & Hemingway, M. (2003). Designing and implementing global staffing systems: Part II—Best practices. *Human Resource Management*, *42*, 85–94.

Salgado, J. E., & Anderson, N. (2001). Cognitive and GMA testing in the European Community: Issues and evidence, *Human Performance, 15*, 75-96.

Sandal, G. M., van, d. V., Bye, H. H., Sam, D. L., Amponsah, B., Cakar, N., . . . Tien-Lun Sun, C. (2014). Intended self-presentation tactics in job interviews: A 10-country study. *Journal of Cross-Cultural Psychology, 45*(6), 939-958. doi:http://dx.doi.org/10.1177/0022022114532353

Schinkel, S., van Dierendonck, D., van Vianen, A., & Ryan, A. M. (2011). Applicant reactions to rejection: Feedback, fairness, and attributional style effects. *Journal of Personnel Psychology, 10*(4), 146-156.

Shackleton, V., & Newell, S. (1994). European management selection methods: A comparison of five countries. *International Journal of Selection and Assessment, 2*, 91-102.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA: Sage.

Steiner, D. D. (2012). Personnel Selection across the Globe. In N. Schmitt (Ed). *The Oxford Handbook of Personnel Assessment and Selection* (pp. 740-767). New York: Oxford University Press.

Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United State. *Journal of Applied Psychology, 81*, 134-141.

Steiner, D. D. & Gilliland, S. W. (2001). Procedural justice in personnel selection: International and cross-cultural perspectives. *International Journal of Selection and Assessment, 9*, 1-14.

Sully de Luque, M., & Javidan, M. (2004). Uncertainty Avoidance. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 602-653). Thousand Oaks, CA: Sage.

Tippins, N.T. (2015). Technology and assessment in selection. *Annual Review of Organizational Psychology and Organizational Behavior, 2,* 551-582.

Wright, P. M., Gardner, T. M., Moynihan, L. M., Park, H. J., Gerhart, B., & Delery, J. E. (2001). Measurement error in research on human resources and firm performance: Additional data and suggestions for future research. *Personnel Psychology, 54*, 875-901.