Abstract. A review of nearly three decades of cross-cultural research shows that this domain still has to address several issues regarding the biases of data collection and sampling methods, the lack of clear and consensual definitions of constructs and variables, and measurement invariance issues that seriously limit the comparability of results across cultures. Indeed, a large majority of the existing studies are still based on the anthropological model, which compares two cultures and mainly uses convenience samples of university students. This paper stresses the need to incorporate a larger variety of regions and cultures in the research designs, the necessity to theorize and identify a larger set of variables in order to describe a human environment, and the importance of overcoming methodological weaknesses to improve the comparability of measurement results. Cross-cultural psychology is at the next crossroads in its development, and researchers can certainly make major contributions to this domain if they can address these weaknesses and challenges.

Keywords: cross-cultural psychology, personality assessment, culture

People have long been interested in other cultures, and contact between people of different cultures and ethnic groups is not a new phenomenon. Throughout human history, people have traveled all over the world for different reasons. As early as the first century of the Common Era, the Roman historian Gaius Cornelius Tacitus (ca. 56–ca. 117 CE) described in his travel notes how the German tribes lived far to the North, providing a very detailed description of their physical characteristics, lands, laws, and customs. He also commented on the general character of the Germanic peoples, particularly on their fierce and independent spirit. In doing so, however, Tacitus was also speaking of Rome in comparison, commenting as much on the Rome of his own time as on the German tribes.

The first deliberate attempt to collect cross-cultural data was probably a missionary model. Hugo Magnus (1842–1907), for instance, an ophthalmologist working in Breslau, was interested in the claim that people from “primitive” nations may be color-blind, something four-time British Prime Minister William Gladstone had once maintained about Greeks living at the time when Homeric poetry was created (Deutscher, 2010). He sent color probes to missionaries and merchants around the world who reported back from the most remote corners of the earth, all indicating that the savages’ color vision was pretty similar to that of people living in a more “enlightened” conditions. Of course, this method was sometimes the only one available, to find answers to questions they had struggled with.

The second, an anthropological method of data collection, was probably devised by British psychiatrist W. H. R. Rivers (1864–1922) – whom Claude Lévi-Strauss called the Galileo of anthropology – during what is known as the Torres Straits Expedition (Deutscher, 2010). The advantages of this more sophisticated method over one using nearly anecdotal missionary reports are obvious: It showed that it was possible to bring experimental laboratory methods even to places that were untouched by civilization. The only serious limitation was low productivity: Well-prepared and well-equipped expeditions could cover only a limited number of geographic locations.

The third method of cross-cultural data collection is unrepeatable nowadays, at least in its original form. To celebrate the 100th anniversary of the Louisiana Purchase, The World’s Fair was held in St. Louis, Missouri, in 1904. Max Weber, who was amazed by everything that American capitalist spirit had to offer, especially by the dazzling lights of the Palace of Electricity, was among the many visitors (Fergusson, 2011, p. 260). However, the main attractions were “savages” who had been brought to St. Louis from all over the world and arranged into “villages” ordered according to their supposed closeness to the animal kingdom (Deutscher, 2010). Robert Woodworth (1869–1962), one of the founding fathers of American psychology, was prepared to receive this unexpected gift by carrying out an extensive psychological study. In his report published in Science, he concluded that


Table 1
The year’s most often cited empirical study of studies published in the Journal of Cross-Cultural Psychology from 2000 to 2009

<table>
<thead>
<tr>
<th>Article</th>
<th>Sample size</th>
<th>No. of cultures or groups</th>
<th>No. of citations (May 1, 2011)</th>
</tr>
</thead>
</table>

“we are probably justified in inferring . . . that the sensory and motor processes, and the elementary brain activities, though differing in degree from one individual to another, are about the same from one race to another” (Woodworth, 1910, p. 179). A century later, in the same journal, Jones (2010) claimed that psychologists are committing a grave error by studying mostly *weird* subjects – people from Western, educated, industrialized, rich, and democratic cultures – who are some of the most psychologically unusual people on the planet. Indeed, a relatively recent analysis of papers published in the top journals of six subdisciplines of psychology between 2003–2007 showed that 96% of the subjects tested came either from the United States or other English-speaking and/or European countries. Only 3% came from Asia, 1% from Latin America, and less than 1% from Africa or the Middle East (Arnett, 2008). This perfectly illustrates that studies of psychological phenomena should also take environmental or cultural variables into account, including populations from a variety of regions across the world. Thus, it is an urgent and crucial matter for researchers to consider and include environmental and cultural heterogeneity in their work.

Many contemporary cross-cultural psychologists have adopted the anthropologist research model. Most typically, researchers choose an “exotic” culture and compare students from that culture to students from their home university, which is located somewhere in North America or Western Europe. Nowadays, it is not even necessary to organize expeditions since many of these exotic places have university professors who can conduct the required studies. In order to more closely observe the standard practices of today’s cross-cultural researchers, we took a closer look at the premier mouthpiece of the community, the *Journal of Cross-Cultural Psychology*. For example, there were 60 articles published in 2009, 45 of which reported data from original investigations that had been subjected to a quantitative analysis. The remaining 15 papers were editorials or meta- or qualitative analyses. The above-mentioned 45 papers reported 58 separate studies with a total of 18,143 participants. Thus, on average, there were about 312.8 participants per study (*Mdn* = 185). These are very impressive numbers, especially in view of the fact that several studies did not merely use paper-and-pencil questionnaires, but were true experiments. In comparison, a relatively recent meta-analysis of temporal instability of personality trait measures demonstrated that the mean sample size across the 75 articles analyzed was only 88.9 participants (*Mdn* = 67) (Watson, 2004). Thus, the standard practices of cross-cultural researchers favorably differ from those of person-
ality researchers who seem to have adopted unacceptably low standards with respect to the size of their samples. However, of the 58 separate studies published in the Journal of Cross-Cultural Psychology in 2009, only 15 (26%) involved a comparison of more than two different ethnic, racial, or cultural groups. In over half of these studies (52%), only two groups were compared with respect to some psychological phenomenon. Most frequently, as one might guess, American students were compared with their peers in an East-Asian country like China, Japan, or Korea. As strange as it may sound, there were even 13 studies (22%) in which all participants were from the same culture.

The anthropological two-culture model has not only been the most popular, but is also quite influential. Table 1 shows most often cited empirical study from the respective year among those published in the Journal of Cross-Cultural Psychology during the 10-year period from 2000 to 2009. Although the studies with a sample size of over 10,000 and participants from up to 56 cultures or territories had the greatest impact, studies contrasting only two cultures or ethnic groups were also able to reach the top of the citation rankings.

The popularity of the two-culture research model seems to be endorsed by theoretical arguments as well. Nearly three decades of cross-cultural research has been dominated by the concepts of individualism and collectivism, which were triggered by Geert Hofstede’s (1980) landmark study of work-related values in the 40 IBM national subsidiaries (Hofstede, 1980). Unlike social scientists in general, psychologists transformed the juxtaposition of individualism and collectivism, which served as a contrast between modernity and traditional authority, into a narrower juxtaposition of the so-called individualistic West and the collectivistic East. In other words, in psychology, individualism is related not only to the idea of modernization, but also to the concept of Westernism or Americanization: If a certain country or cultural group (e.g., East Asian, former Communist countries) has not fully adopted the Western/American program of modernity, it is most likely to be considered collectivist (Allik & Realo, 2004). This conceptualization, in turn, provides a certain justification for the adequacy of two-culture comparisons, one culture typically from the affluent West (United States or Canada), the other from East Asia, typically represented by China (most often Hong Kong), Japan, or Korea. In some cases, this conceptualization serves as a good excuse for avoiding more informative but less accessible comparisons to, say, African cultures (Rossier, Dahourou, & McCrae, 2005; Rossier, Rigotti, & Personality Across Culture Research Group, 2008) or cultures populating other non-Western countries such as the Russian Federation (Allik et al., 2009, 2011). In fact, assessing cultural differences more accurately and with a larger set of indicators is one of the major goals we must strive to achieve to significantly improve cross-cultural research in the near future (Duarte & Rossier, 2008).

Although cross-cultural researchers are habitually overoptimistic about developments in their own field (van de Vijver, 2006), there are obvious signs of exhaustion. For instance, it is hard to deny that the most influential recent advancements in cross-cultural psychology were made by researchers who may not consider themselves cross-cultural psychologists. There is no doubt that even collecting data from only two cultures is rather expensive and time-consuming, let alone from a few dozen. There are only a few effective ways to collect cross-cultural data from a sufficient number of countries to truly improve cross-cultural research. The first is to develop a popular questionnaire that is then translated into a large number of languages by enthusiastic colleagues. The Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975) and the NEO PI-R are good examples of this relatively slow and inefficient method of data collection (Lynn & Martin, 1995; McCrae, 2002; van Hemert, van de Vijver, Porzinga, & Georgas, 2002). Compared to these achievements, it is truly astonishing that cross-cultural researchers have not been able to agree upon the nature or number of attributes that are essential for defining and measuring individualism and collectivism (Realo & Allik, 2008; Realo, Koido, Ceulemans, & Allik, 2002). After 30 years of enthusiastic praise, devastating criticism, and thousands of articles written on individualism-collectivism, there is still no generally accepted instrument that could be translated into as many languages as possible. This clearly suggests that it is truly insufficient to consider only individualism-collectivism when characterizing a culture. Rather, we need to theorize and identify a larger set of variables in order to describe a human environment. Hofstede’s (1980) work is certainly a valuable starting point, but more effort should be made in this domain, as illustrated by the paper by Fischer (this issue).

Another way to carry out large-scale cross-cultural comparisons is to form an international research consortium held together by the promise that the first two or three papers are coauthored by everyone who participated in collecting data. Again, these were evolutionary psychologists, not so-called “genuine cross-cultural researchers,” who developed such a powerful research method (Buss et al., 1990), allowing, for example, for the collection of personality data from over 50 countries or territories (Schmitt & Allik, 2005; Schmitt, Allik, McCrae, & Benet-Martinez, 2007). Using the same method, McCrae and Terracciano collected not only self-reports, but also observer-reported data from 50 cultures (McCrae, Terracciano, & 78 members of the Personality Profiles of Cultures Project, 2005; McCrae, Terracciano, & 79 members of the Personality Profiles of Cultures Project, 2005).

In a typical two-culture study, not only the selection of countries, but also the selection of participants is based on convenience rather than representativeness. For many years, comparative research projects in psychology (including the above-mentioned large-scale personality studies) were conducted using mostly student populations. More recently, more and more researchers interested in cultural differences turned their attention to large-scale social studies, such as the World Values Survey (WVS), which in collaboration with the European Values Study (EVS) managed to collect data from representative national samples.
in 97 societies containing almost 90% of the world’s population (World Values Survey 1981–2008 Official Aggregate V, 2009), and the European Social Survey (ESS), which conducted five rounds of surveys covering over 30 nations and employing the most rigorous research methodology (e.g., Davidov, Schmidt, & Schwartz, 2008). Instead of vague comparisons between two more or less arbitrarily chosen cultures, these studies provided comprehensive maps showing how people around the world think, feel, and behave (Inglehart, Basanez, Diez-Medrano, Halman, & Luijkx, 2004).

EAPP Expert Meeting on “Personality and Culture”: A Look Forward

With the support of the European Association of Personality Psychology, an expert meeting on the topic of “Personality and Culture” was held at the University of Lausanne in Switzerland from September 29 to October 1, 2010. The meeting was organized by Jérôme Rossier and Koorosh Massoudi from the University of Lausanne, Switzerland, and by Anu Realo and Jüri Allik from the University of Tartu, Estonia. During the discussions, participants repeatedly demonstrated that cross-cultural psychology has reached the next crossroads. Although studies contrasting only two cultures may occasionally provide some useful information, multicultural studies involving many different countries have shown that there is a considerable amount of information beyond comparisons of the so-called individualistic West and the collectivistic East (Proyer et al., 2009; Schmitt & 118 members of the International Sexuality Description Project, 2003). These large-scale cross-cultural studies have demonstrated that the traits studied have a meaningful geographical distribution (Allik & McCrae, 2004; Inglehart et al., 2004; Rentfrow, Gosling, & Potter, 2008; Schmitt et al., 2007). However, the geographical gradients do not always make perfect sense. The rankings of cultures on some personality traits may seem staggering unexpected. Indeed, it is not highly predictable that the most conscientious – determined, strong-willed, organized, dutiful, and deliberate – people live in Burkina Faso and Congo, while the least conscientious (at least according to self-reports) live in Japan and Korea (Mõttus, Allik, & Realo, 2010). Apparently in support of these casual observations, Heine, Buchtel, and Norenzayan (2008) reanalyzed published data and showed that aggregate national scores for self-reported conscientiousness were – contrary to the authors’ expectations – negatively correlated with various country-level behavioral and demographic indicators of conscientiousness, such as postal workers’ speed, accuracy of clocks in public banks, accumulated economic wealth, and life expectancy at birth. Oishi and Roth (2009) extended the list of paradoxical findings by showing that nations with high self-reported conscientiousness were not less but more corrupt.

Among the attempts to resolve the apparent or real puzzle, Leon Festinger’s social comparison processes – the idea that people estimate their attitudes or dispositions relative to social standards (Festinger, 1954) – has repeatedly been revitalized. For example, Heine and colleagues (2008) proposed that people likely bring to mind a standard that lies outside their own culture, for example, a perceived international norm (Heine et al., 2008), when providing self-reports. Even if this explanation seems almost irresistible, it is based on the unrealistic assumption that a layperson has the extraordinary ability to obtain accurate information about mean levels of personality traits across many countries. In spite of the fact that these frame-of-reference explanations look deceptively simple, nobody has attempted to measure these hypothetical social standards. In order to ameliorate this abnormal situation, a cross-cultural study was carried out to test whether people from 21 countries (Australia, Benin, Burkina Faso, the People’s Republic of China, Estonia, Germany, Hong Kong, Japan, Lithuania, Malaysia, Mali, Mauritius, the Philippines, Poland, Russia, Senegal, South Africa, South Korea, Sweden, Switzerland, and the United States) have different standards for conscientiousness (Mõttus et al., 2011). All participants rated their own level of conscientiousness and that of 30 hypothetical individuals portrayed in short vignettes, whereby the latter ratings were expected to reveal individual differences in standards for conscientiousness. Contrary to the expectations of reference-level theorists, the vignettes were rated relatively similarly in all cultures, suggesting no substantial culture-related differences in standards for conscientiousness. Controlling for the small differences with respect to these standards did not substantially change the ranking of countries on mean self-ratings or the predictive validity of these rankings for objective criteria. These findings lend little support to the hypothesis that mean self-rated conscientiousness scores are influenced by culture-specific standards, considerably restricting the range of potential explanations for the puzzling country rankings for conscientiousness (Mõttus et al., 2011). Although it is premature to draw any firm conclusions, personality traits may be estimated in absolute rather than relative terms. For example, people may have developed a more robust and unconditional way of judging their basic tendency to feel, think, and behave than of judging the level of political freedom in their society or their work satisfaction, which typically show the reference effects of the social standards (King, Murray, Salomon, & Tiandon, 2003; Kristensen & Johansson, 2008).

There seems to be consensus that the most pressing problem of cross-cultural comparisons is the comparability of measurement results. Although methodological activists have repeatedly urged others to test measurement invariance across translated instruments and cultures (Byrne & Watkins, 2003; Little, 2000; Vandenberg & Lance, 2000), practitioners have been rather reluctant to do so. Although
it is generally acknowledged that the establishment of measurement invariance across groups is a logical prerequisite for comparing data collected from two groups in two different locations, in practice measurement invariance is rarely tested in practical cross-cultural studies. The reluctance of practitioners to adopt new practices is understandable: Most measurement invariance applications are based on the item response theory (IRT), which assumes that all individual responses must be congruent with the general response pattern of a group. Unlike intelligence tests, there is no one common response pattern in personality and attitudes since we expect some individuals to score high and others low on personality trait or attitude measures. Indeed, the assumption that some personality items represent more extreme expressions of a trait than others (the so-called “item difficulty”) and that this pattern of endorsement needs to be identical across individuals is a premise that still needs to be proven (Alik et al., this issue).

The second reason for the unpopularity of the measurement invariance agenda is the failure of methodologists to explain to practitioners that the testing of invariance is not a control, but rather an amelioration issue. First of all, measurement invariance is not a binary category. There are several degrees of invariance, and the toughest metric invariance is not even needed to solve many problems (Davidov, 2000, 2008, 2009). Very often, the testing of measurement invariance on collected data is done too late. Cross-cultural researchers are starting to realize that it is wiser to test measurement invariance during the test construction and translation phase. Indeed, some of the already adapted instruments may need a revision that entails substituting a number of items. All instruments require revision sooner or later, and a lot of time and money would be saved if these revisions seriously considered measurement invariance issues. Recently, the validation of a specific language version of a psychological measurement instrument adopted a multicenter cross-national approach to study the same instrument in four heterogeneous samples (viz., from Belgium, Canada, France, and Switzerland) in terms of age and sex. This study clearly indicated that a high level of measurement invariance can be reached if one uses the same language version of the instrument in four very similar cultures, even if they differ on the uncertainty avoidance dimension with higher scores in the French-speaking part of Belgium and in France and lower scores in the French-speaking parts of Canada and Switzerland (Rossier, Hansen, Baudin, & Morizot, in press). This study indicated that a particular language version can be used in different countries or cultures with the same norms. However, the study also suggested, as stated above, that more studies should adopt a multicentric cross-national approach (Rossier, 2005).

In conclusion, cross-cultural research has considerably matured during the last few decades. As a result of this research, we have reliable information of how more widely (e.g., neuroticism) or narrowly (e.g., fear of being laughed at) psychological traits are distributed around the world, in which there are numerous countries or territories. With the expansion of research geography, methodology has also improved in the pursuit of making results more reliable for theoretical and practical generalizations.

Current Issue

The current issue represents a selection of contributions based on some of the most stimulating studies presented at the European Association of Personality Psychology Expert Meeting held at the University of Lausanne. We would like to acknowledge the support of the Swiss Psychological Society and the editors of the Swiss Journal of Psychology, who have allowed this special issue on culture and personality to emerge. The four contributions concern very different topics, use different methodologies, and were written by colleagues from a variety of countries.

In his contribution “Intersubjective Culture: Indeed Intersubjective or Yet Another Form of Subjective Assessment?” Ronald Fischer presents a review of the potential definitions of culture. Indeed, one’s definition of culture has implications for the methodology one implements to measure it. This review of the different existing conceptualizations of cultural environments has serious implications for the most central constructs used by cross-cultural psychologists and thus represents a very interesting and challenging starting point for this issue. Ronald Fischer argues that an intersubjective consensus approach may allow us to identify important values or issues that characterize a cultural or ethnic group. We should surely admit that cross-cultural researchers have been rather insensitive to the type of instructions used to collect people’s ratings or opinions. Sometimes they even naively believe that asking how important values or other issues are for people in one’s cultural or ethnic group (or country) is essentially the same as asking for self-reports (see McCrae, Terracciano, Realo, & Allik, 2008). Ronald Fischer presents an interesting study in which, instead of using self-reports about what is important for oneself, the researchers ask respondents to rate how important each value or issue is for people in one’s cultural or ethnic group or country. An analysis of intersubjective norms, however, demonstrated that consensus for these norms was low, and that relatively little variance was explained by cultural background. According to the author, this challenges the intersubjective norm approach as a construct, which is able to capture culturally shared meaning systems. Indeed, culture might also be defined as an aggregate of heterogeneous interconnected environments. If so, more attention should certainly be devoted to better assessing individuals’ subjective perception of their environment and to studying these perceptions along with other individual characteristics or processes.

In “A Review of the Impact of Personality and Culture on the Job-Demands-Control Model of Job Stress,” Christina Györkös, Jurgen Becker, Koorosh Massoudi, Gideon
de Bruin, and Jérôme Rossier review the psychological scientific literature on job-related stress. The job demand-control model is certainly one of the most influential approaches to the measurement of work strain, allowing us to empirically study the outcomes of work stress, and to the identification of variables that might moderate the relationship between work strain and work stress. Research on personality and work stress has clearly shown that some personality dimensions (e.g., neuroticism) increase one’s vulnerability to work strain, whereas other dimensions (e.g., conscientiousness) might be a protective factor in the stress process (Massoudi, 2009). The current trend toward global integration of business and the disappearance of geographical boundaries in many organizations tend to increase worldwide competitiveness and cultural diversity at the workplace. Thus, it seems important to study the extent to which the cultural environment and people’s cultural values might have an impact on their resilience or vulnerability when coping with job stress. Moreover, frequent relocations might have a direct impact on an individual’s social network and thus affect the potential social support they can benefit from, as social support is a well-documented protective factor against job stress. This review clearly suggests that more studies are needed to describe in greater detail the combined impact of stable individual characteristics, such as personality dimensions and environmental or cultural factors on the reaction to job stress.

We find it quite interesting that the concept of morality has started to regain its former importance in psychology. In modern evolutionary psychology, altruistic behavior is seen as crucial to group cohesion and consequently to the survival of the species. In their article “Are Virtues Shaped by National Cultures or Religions?” Jan Pieter van Oudenhoven, Boele de Raad, Carmen Carmona, and Meta van der Linden studied the impact of religion and national cultures on the value of virtues, which are positive or “morally good personal characteristics,” in four different religious groups in the Netherlands and in three other European countries. The question was to study whether the impact of religion on the relative value of virtues was more important than the influence of national cultures. They observed that the rating of virtues did not differ very much with respect to gender or age. Differences across religions were also very small. The largest differences were observed across countries, even if they were all European, indicating that national culture seems to have a larger impact on the relative intersubjective value of virtues. The differences were very heterogeneous with almost no difference for most virtues and rather large differences with respect to reliability and openness. Even though the results of this study are very interesting, it would be useful, as stated by the authors themselves, to extend this study to other moral ideas to include for example, religious and moral values. Five Factor Model (FFM), a very well-validated model for capturing interindividual differences in personality at the individual level. The authors’ approach is innovative and their conclusions are in line with the intuition of most personality psychologists, but nevertheless contradict the conclusions of researchers who claim that inter- and intraindividual differences cannot be captured using the same set of variables. The problem with such a view is that if you study measurement differences of a stable trait, you are basically studying measurement error or noise. Allik and colleagues studied the structure of covariance at the individual level using two person-fit indices with participants from four European countries. They demonstrated that the double-entry intraclass correlation (ICC DE) of the 30 NEO PI-R scores is a good person-fit index that demonstrates how adequate the FFM of personality is for each individual. As for the 5% of subjects for whom the structure of covariance is inconsistent, their heterogeneity within a dimension might be considered a specificity characterizing their profile. Of course, it would be very interesting to study in more detail the behaviors and feelings of such individuals who present an atypical combination of personality traits. These results clearly indicate that variable-centered and person-centered approaches are consistent.

Conclusion

We hope you will find many new and interesting insights in this special issue of the Swiss Journal of Psychology on personality and culture. We wish you pleasant reading and hope that this issue will stimulate your own thoughts and encourage you to take environmental and cultural variables into account more systematically when studying psychological phenomena.

Acknowledgments

The writing of this article was supported by a grant from the Estonian Ministry of Science and Education (SF0180029s08) to Jüri Allik and by a Swiss South African Joint Research Program grant from the State Secretariat for Education and Research (SER) to Koorosh Massoudi and Jérôme Rossier (SSAJRP no. 11).

References
