ORIGINAL RESEARCH PAPER

Volume-8 | Issue-3 | March-2019 | PRINT ISSN No 2277 - 8179

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

SOLUTIONS FOR THEORETICAL AND METHODOLOGICAL CHALLENGES OF COMPARATIVE CRIMINOLOGY RESEARCH

Social Science

Myunghoon Roh

Department of Social Science and Humanities, Texas A&M University – San Antonio
One University Way, San Antonio, TX 78224

KEYWORDS

This study is comprised of two parts. The first part explores more closely some theoretical and methodological challenges of the comparative method in criminology. From a theoretical perspective, there is a need to develop and test crime theory that is applicable across different national and cultural settings. From a methodological perspective, there is a need to properly establish measurement invariance that is a prerequisite for valid cross-national comparisons. Finally, in the second part, the current study suggests some solutions for theoretical and methodological challenges of comparative criminology research.

Comparative Research: theoretical and methodological challenges

Although comparative research has advanced our understanding of national cultural aspects of social phenomena, it has both challenges theoretically and methodologically. This part discusses some of these challenges. Comparative research typically attempts to explain social phenomena using theories developed within one country (Farrington, 2000). Given that most theories were originally developed for use in one particular culture, researchers cannot readily draw from universal or broad existing theories to export ideas developed in one cultural setting to another setting (Karstedt, 2001). Theories developed within single cultural settings may be too simplistic or idiosyncratic to explain social phenomena in a variety of different settings. The main benefit of comparative research is that it allows for the assessment of the generalizability of certain theoretical frameworks. Thus, one challenge is to identify a theoretical framework which has the potential to be generalizable across different countries.

The primary methodological challenge for the comparative method is translating a concept from one cultural context into another one. Øyen (1990), a social anthropologist, discussed the methodological challenges related to observations drawn from within the Western culture and applying them to other cultures. Specifically, distortion of the content and meaning of the concepts must not be permitted, and valuable information must not be lost through translation (Øyen, 1990). Thus, scholars in social sciences have mainly focused on conceptual equivalence, an equivalence of measurement, and sampling (Warwick & Osherson, 1973).

Conceptual equivalence is the basic theoretical question of whether concepts under study have equivalent meanings across the social unit understudied (Warwick & Osherson, 1973). For conceptual equivalence, three main criteria have to be established. The first one is the dimension of universality. Warwick and Osherson (1973) identified general notions in Western cultures (e.g. mother, incest, socialization, illness, deferral bureaucracy, and arctic hysteria) and claimed that these concepts are unlikely to be universally understood in all cultures. Definitional comparability is the second aspect of conceptual equivalence. It holds that concepts do not have to be necessarily identical in meaning, but conceptual equivalents have to exist. Smelser (1976), a comparative researcher, highlighted the difficulties of defining economic concepts in particular. In addition, Leighton and colleagues (1963) pointed out the importance of comparable definitions of mental disorders in cross-cultural studies. They found many psychiatric definitions derived from Western cultures (e.g. mental deficiency, psychologically expressed disorders, and psychologically derived disorders) are not necessarily applicable to other societies. The final aspect of conceptual equivalence is the identifiability of concepts. This is the assumption of whether or not any concept can be applied to any given cultural settings. For instance, while the concept of a "mother" is readily identified observationally in almost every culture, there is no analogous term for "mental deficiency" across cultures. Warwick and Osherson (1973) emphasized that wide variation in the identifiability of a central concept results in methodological challenges for comparative research, including formidable problems of measurement. Even if comparable conceptual definitions in the societies under study exist, comparative researchers further face the challenge of developing equivalent indicators (Warwick & Osherson, 1973). Ultimately, these issues suggest the overall difficulties of transitioning from theoretically meaningful concepts to empirically observable manifestations.

Comparative researchers have identified five overlapping problems related to an equivalence of measurement: differential researchability, comparability of stimuli, context, response, and reliability-validity (Warwick & Osherson, 1973). The first aspect of equivalence of measurement, research ability, needs to be considered because theoretically applicable concepts may differ in different cultures under study. For instance, the notion of "looking for work" (a key element in studies of unemployment in the developed world) is irrelevant to tribesmen and peasants in many parts of the Third World (Warwick & Osherson, 1973). In addition, the unwillingness of respondents to answer questions on sensitive topics is a barrier to research ability. Respondents may feel uncomfortable participating in comparative research because of the perceived harm it could do to their community and society. Another example of research ability is when respondents are unable or unaccustomed to discussing a certain topic. As an example, in the study of the socialization of adolescents, Danish teachers claimed that they were unaccustomed to analyzing matters of adolescents' interactions and could express no judgment or opinions on the subject (Stodolsky & Lesser, 1973).

When conducting comparative research, different indicators are often needed to capture the same concept in different cultures. Intelligence can be defined as the ability to adapt effectively to the society in which one lives (Warwick & Osherson, 1973). However, the appropriate indices of adaptation will be different in rural Asia as it is in the urban United States. Thus, for comparability, all items must be culture-specific. There are several ways to approach the question of equivalence in stimuli. Before comparative research is conducted, the collaboration between knowledgeable members of all the participating societies must be consulted. If collaboration is not possible, some type of preliminary exploration has to be investigated through ethnographic reports, travelers' accounts, and the like (Warwick & Osherson, 1973). According to Warwick and Osherson (1973), social scientists can never guarantee that a set of questions developed in one society can be translated and exported for use in another cultural setting. If social scientists from different countries in a certain comparative study are not involved in choosing the concepts and developing the questions, they cannot be sure the items in the study are formally identical and functionally equivalent in meaning from society to society. At the very end of the study, it is often difficult to determine whether different findings reported in the study stem from variations in national cultures, from different meanings attached to the questions in the countries, or from other irrelevant factors in the study.

Based on the methodological challenges mentioned above, when using comparative research methods empirically, researchers encounter three related methodological problems, all of which can be categorized under the problem of comparability. The first problem is whether the dependent variables, such as the events and situations the researcher wish to explain, are comparable from one sociocultural context to

another (Warwick & Osherson, 1973). For instance, one could ask how we can compare crime rates among different countries using data collected centuries ago. The second problem is related to the general dimensions used to compare societies, whether the dimensions distort events and situations cross-culturally. If the political, economic, social, or cultural aspects of the two societies are compared, we are not sure whether these dimensions are comparable in both societies. The third problem is when different social units are used for comparison. It is not appropriate to compare societies that have very different social systems, such as between a highly complex nation-state (the United States) and a small tribe in an Asian country.

Theoretical integration as a way to overcome cross-national theoretical challenges

As already alluded to briefly in the previous section, theoretical integration has been suggested as a way to overcome the challenges related to theory development and testing within a cross-national context. Some researchers have suggested that integrating existing theories into a comprehensive framework may be one potential solution for the theoretical poverty of comparative research (Pratt & Godsey, 2002). In addition, more recently, researchers emphasize the need to include macro-level social change in their explanatory frameworks, as well as suggesting micro-macro level integration as a new avenue of theoretical development. Such developments would account for diversity in national contexts as well as historical change (Messner, 2011).

Theoretical integration is, generally, defined as "the act of combining two or more sets of logically interrelated propositions into one larger set of interrelated propositions, in order to provide a more comprehensive explanation of a particular phenomenon" (Klein, 1989, p. 75). According to Muftic (2009), there are generally three goals of theoretical integration; theory reduction, increasing explained variance, and theory development through the clarification and expansion of existing propositions and theoretical concepts. Although opponents of theoretical integration argue that integration is not possible because of differences in underlying philosophical assumptions and the complexity of integrated theoretical models (Hirschi, 1969, 1989), proponents, however, argue that theory competition can inhibit theoretical development (Bernard & Snipes, 1996). By integrating existing theories, researchers explain a greater portion of the variance that remains unexplained by separate theories (Messner et al., 1989).

Especially, the micro – to – macro linkage is one of the most promising areas of future research for the solutions of theoretical challenges of comparative research. By integrating micro- and macro- level variables, comparative researchers can recognize the possible role of different levels of determinants, such as social structure, on an individual's behavior (Messner, 2011).

Macro-level theories "link social structure characteristics to variations in the rates and distributions of crime" (Bernard, Snipes, & Gerould, 2016). On the other hand, micro-level theories "link individual characteristics to the probability that an individual will engage in criminal behavior (Bernard, Snipes, & Gerould, 2016). Macro-micro theoretical integration combines macro- and micro-level theoretical explanations. The principal weakness of macro-level theories is their inattention to "personal motivation or the agency of the individual offender (Barak, 2008, p. 197). On the other hand, the primary weakness of micro-level theories is their inattention to "the context within which individuals are embedded and, more specifically, the vulnerability of micro-level processes to local economic and social conditions" (Bellair, Roscigno, & McNulty, 2003, p. 25). Thus, integrated macro-micro theories examine the effect social structure has on individual characteristics and subsequent individual action (Paternoster & Bachman, 2001). Identifying mediating and moderating linkages between macro- and micro-level variables, macro-micro theoretical integration differentiates the causal properties of structural and individual factors.

A fundamental assumption underlying the micro – to – macro linkage is that individuals' social behaviors are determined to some extent by social forces in their wider environment. The importance of linking both theoretical frameworks examining both the macro and micro levels into comparative research can be justified in several ways. Most sociological theories assume that social structural conditions have a direct impact on an individuals' behavior independent of their personal characteristics. Thus, if comparative studies exclude this wider context, they would suffer from serious problems with model specification. Second, many individuals – level effects are actually reflective of neighborhood, community, and country – level dynamics. For example, the strong impact of the financial situation of a certain country on an individual's economic behaviors is commonly attributed to the lifestyle of such persons. Under this condition, if we exclude the contextual factor, this would not specify the true relationship between an individual's characteristics and economic behavior. Thus, macro measures are important in comparative research, because they may directly influence human behavior and explain the effects of individual-level variables.

Measurement Invariance Testing as a way to overcome cross-national methodological challenges

In comparative research, the same instruments are generally used for all involved groups of interests. Researchers who conducted this kind of comparative research frequently assume that obtained results are comparable for valid comparison across the groups. However, this assumption of results' comparability is untested and focuses only on the difference in average scores of the two or more cultural groups. Therefore, when testing for cross-cultural differences in any comparative research, a fundamental concern of ensuring invariance arises (Hui & Triandis, 1985). Therefore, without strong evidence of invariance, comparisons of the survey scores across different countries may be potentially erroneous or misleading.

As claimed by comparative scholars, the lack of evidence for measurement invariance raises concerns about whether the survey items capture the same theoretical construct in other countries or cultural settings (Vandenberg & Lance, 2000). In the end, this concern may raise suspicion that the results of cross-national comparisons might not be meaningful or valid. Evidence of measurement invariance is one of the most important methodological challenges of comparative research. Specifically, differences in observations across different countries might be due to true differences between countries or they may be due to systemic biases related to the different ways of responding to certain items between different countries. Thus, cross-national differences of scale scores might reflect real differences of constructs across countries and differences in scale reliability rather than reflect non-invariance of the constructs measured from all involved countries (Steenkamp & Baumgartner, 1998). Accordingly, researchers should ask: do the survey items offer a cross-nationally valid measurement across different countries? Cross-national researchers have to test and establish measurement invariance in order to determine whether measurements capture the same underlying constructs in each participating country under study.

Measurement invariance refers to an instrument's ability to measure the same construct in the same way across different groups (Millsap & Kwok, 2004). Measurement invariance has been tested by comparing means and variances across different groups. However, comparing means and variances is not the same as comparing the underlying measurement structure across sub-groups (Vandenberg & Lance, 2000). A central premise of measurement invariance is that the scores across groups are on the same scale. In addition, the relationship between test items and theoretical constructs of interest must remain stable across groups (Reise, Widamand, & Pugh, 1993). That is, the measurement relates to the theoretical construct of interest (e.g. attachment, risky routine activities, and self-control) in the same way across different groups. Therefore, researchers should compare whether the number of factors and the items that make up a factor is similar across different groups. Comparing differences in means is required to determine if there is an equivalent underlying model. Thus, it is essential to investigate the means and measurement structure for valid comparison across different groups (Gregorich, 2006).

When a measurement produces the same observed scores for individuals who have identical attributes being measured, measurement invariance is established (Meade & Bauer, 2007; Schmitt & Kuljanin, 2008). Measurement invariance testing has generally employed several measurement conditions (Meade & Bauer, 2007; Vandenberg & Lance, 2000). The generalizability of scale or measurement across subgroups of a population is the first type. Sometimes, it may be required to generalize among subgroups which are located in a single culture or diverse cultures. For instance, when research is conducted among different gender, age, race, ethnicity, area, country, and subcultural groups, it is required to establish

invariance testing to see whether a scale or measurement developed in one culture works in other ones as well (Byrne et al., 1989; Rensvold & Cheung, 1998). Whether a test or measurement developed in a certain country may be translated and utilized in a different country is an important methodological question for cross-national research method (Cheng & Rensvold, 2000). Thus, measurement invariance testing is of great concern for cross-cultural researchers.

According to Yoon and Millsap (2008), measurement invariance is a broad term encompassing both linear and nonlinear relationships between observed variables and latent factors with special attention to the whole score distribution. In a factor model, measurement invariance is called factorial invariance (Meredith, 1993; Yoon & Millsap, 2008), which is expressed within a linear factor model with mean and covariance. Measurement invariance testing is generally conducted with linear confirmatory factor analysis (CFA), which is defined as the equivalence of parameters specified in the model across groups. Thus, there are different levels of factorial invariance, which can be determined depending on the parameters in the testing of invariance. For comparative research methods, the general structural equation models can be expanded to multiple group confirmatory factor analysis by applying a group indicator into CFA.

In the multi-group confirmatory factor analyses model, factorial invariance testing covers the first four null hypotheses which test the invariance of variance-covariance matrices of observed variables, factor loadings, intercepts, and unique variance of observed variables, sequentially. Then the last two null hypotheses are for the equalities of factor variance-covariance and factor means over groups, respectively. This order is the step to test multigroup confirmatory factor analysis. More details follow in the next section regarding procedures for testing measurement invariance. As indicated in the prior section, different levels of factorial invariance are tested, depending on the certain set of parameters which are tested for group equality. Researchers must keep in mind two suggestions about the sequence of measurement invariance testing. It is not necessary or easy to attain in reality the full invariance of all levels. In addition, the sequence of different levels of invariance testing is largely determined by research questions and interest of the study (Vandenberg & Lance, 2000).

Configural invariance Researchers are interested in the configuration of a model across different groups. As the most basic level of measurement invariance, the configural invariance testing gauges whether similar factors are measured across the groups (Widaman & Reise, 1997). It tests whether or not participants from different groups conceptualize the construct in the same way. Thus, at this level of invariance, it is not necessary that the set of observed indicators and the construct have exactly the same strength but that the same set of an item must be associated with the same latent factor in each group. The configure invariance testing model works as a baseline model for other types of invariance tests. Comparing each model to the baseline model, researchers determine whether group invariance can be established or not. Researchers have used the configural invariance as "same simple structure" (Meredith, 1993), "weak factorial invariance" (Horn & McArdle, 1992), "baseline model" (Bagozzi & Edwards, 1998), "equality of factor structure" (Cole & Maxwell, 1985), "conceptual equivalence of measures" (Vandenberg & Lance, 2000), "conceptual equivalence" (Hui & Friandis, 1985), and "construct equivalence" (Van de Vijver & Leung, 1997).

Metric invariance The configural invariance testing does not enable a straightforward comparison of results because it does not indicate whether respondents from different groups assign the same meaning to items. On the other hand, metric invariance is a more stringent form of invariance given that it assumes that the relationship between observed indicators and the concept of the latent factor is equal across groups (Cheung & Rensvold, 2000). In other words, metric invariance implies the equality of the measurement unit or intervals of scale on the latent concept is measured across different groups (Steenkamp & Baumgartner, 1998; Van de Vijver & Leung, 1997). Focusing on inter-group equality of slope parameter, researchers can test whether the factor loadings of each item are equal across different groups. If so, the researcher can be confident that items and questions in the survey are understood in a similar way in different groups. Thus, this invariance test enables the testing of instrument measures of the same latent construct in all of the groups in the study. According to Van de Vijver and Leung (1997), meaningful comparisons of regression coefficients and covariances are possible, when partial metric invariance for all groups in the study is obtained.

Scalar invariance For establishing measurement invariance and a full comparison of groups' scores, it is necessary that the scales of the latent construct have the same origin (Meredith, 1993). The scalar invariance refers to the invariance of intercepts across groups. The establishment of both the metric and scalar invariance is useful when the main research question of group comparison research including cross-national research needs the comparison of mean scores across groups. Metric and scalar invariance are necessary for latent mean comparison across groups. For instance, individuals who have the same score on the latent construct would obtain the same score on the observed variable regardless of their group membership. By constraining the intercepts of items to be the same across groups, the scalar invariance is tested. To compare scores across different groups, this model is the last necessary model to test. All additional tests are optional and meaningful in specific research purposes.

This study seeks to understand partial invariance. Methodologists have noted that invariance of the parameters for all items is not necessary for meaningful comparison (Steenkamp & Baumgartner, 1998). In reality, metric and scalar invariance are unlikely in many situations. Thus, researchers can compromise between full measurement invariance and lack of measurement invariance (Kankaras & Moors, 2010). Comparison across different groups can be made in a valid way if at least two items within one construct are invariant (Kankaras & Moors, 2010). One item within a scale needs to be work as a "marker" item to define the scale of each latent construct. Also, to test the invariance of the marker item, one more item needs to be invariant.

Confirmatory factor analysis is a model testing technique in which a theoretical model is compared with the observed structure in a sample (Milfont, Fischer, & Ronald, 2010). To determine the degree to which the theoretical model is consistent with the empirical data, researchers use goodness-of-fit indices after each level of invariance testing is performed. These indices indicate how well the empirical data fit the proposed theoretical model. Statistical tools, such as LISREL, Amos, EQS, and Stata provide several indices to assess how well a hypothesized model fits the sample data. The likelihood ratio test (referred to as the chi-square or X^2test) between a baseline model and sequentially constrained models is an objective model fit index, which has been traditionally used as a goodness-of-fit statistic in structural equation modeling. However, the literature of measurement invariance has reported that the chi-square test is too sensitive to the size of a sample (Bentler & Bonett, 1980). Thus, it has been recommended to use the likelihood ratio test as a measure of fit instead of a test statistic (Bollen, 1998).

To overcome the limitations of the likelihood ratio test, several fit indices have been developed (Hu & Bentler, 1999; Kaplan, 2000; Mulaik et al., 1989). Referred to as absolute fit indices, these fit indices assess improvement in fit by comparing a target model with a more constrained nested model (Hu & Bentler, 1999). Given that diverse fit indices deal with different aspects of fit, researchers should report multiple fit indices in structural equation modeling studies (Thompson, 2000). The absolute fit indices used to evaluate overall model fit are as follows: the likelihood ratio test, the chi-square to degrees of freedom ratio (X^2/df) (Wheaton, Muthen, Alwin, & Summers, 1977), the root mean square error of approximation (RMSEA) (Steiger & Lind, 1980), and a standardized version of Jőreskog and Sőrbom's (1981) root mean square residual (SRMR).

According to methodologists, a X^2/df ratio of 5:1 or less indicates a good fit (Carmines & McIver, 1981); RMSEA and SRMR values close to .06 and .08 respectively indicate acceptable fit (Hu & Bentler, 1999); and RMSEA values in the range of .08 and .10 indicate medium fit and above .10 indicate poor fit (Browne & Cudeck, 1989). In addition, to test the improvements over competing models, several indices should also be considered, such as, the difference in chi-square between two nested models (X^2 difference test), the comparative fit index (CFI) (Bentler, 1990), the expected cross-validation index (ECVI) (Browne & Cudeck, 1989), and a consistent version of Akaike's (1987) information criterion (CAIC) (Bozdogan, 1987). Results from the X^2 difference test indicate that the model with smaller X^2 has a statistically better fit. However, this difference can be influenced by large samples. Thus, the X^2 difference test has been used for significant improvements in models. Alternatively, CFI values close to .95 indicate acceptable fit (Hu & Bentler, 1999). Lower ECVI and

CAIC values reflect the model with a better fit (Garson, 2008). And MacCallum et al. (1996) suggested that 90% confidence intervals (90% CI) were also reported for both RMSEA and ECVI.

Although several techniques have been developed to test measurement invariance (Hui & Triandis, 1985), the multi-group confirmatory factor analysis model has been accepted as the most powerful tool for testing cross-national measurement invariance (Steenkamp & Baumgartner, 1998). Assessing the cross-national comparability of the measurement of criminal behavior and individuals' psychometric measures is rarely done for a number of reasons: the lack of measurement invariance testing literature, the lack of widely unified terminology of measurement invariance testing, criminologists' uncertainty regarding measurement invariance testing for cross-national comparison, and the absence of clear guidelines to test measurement invariance.

Conclusion

The current study on the comparative criminology research suggests ways to deal with current theoretical and methodological challenges of comparative research. First of all, given that most of criminology theoretical frameworks have been developed in American and western cultural settings, comparative criminologists need to test whether it is reasonable to export ideas developed in the United States to other countries. We also want to take into consideration that individual social behaviors are determined by social structural factors in their wider environment. There is a need to integrate macro- and micro-level variables in order to investigate how social structural conditions may have an impact on individuals' behavior and attitudes. In addition, comparability of instruments has to be tested to make sure that the theoretical concepts under study have the equivalent or same meaning across different countries. This may be done by measurement invariance testing, using multiple group confirmatory factor analysis. By testing measurement invariance, configural invariance, metric invariance, and scalar invariance, researchers can evaluate whether similar item factors, the relationship between observed indicators and the concept of the latent factor, and the scale of the latent construct are equal across different groups (Cheung & Rensvold, 2000). In the current study, I have laid the groundwork for the two main prerequisites of optimal comparative criminology research. First, the measurement invariability of the major theoretical concepts should be established. Then, the generalizability of an integrated theoretical framework should also be tested across different countries.

REFERENCES

1. Akaike, H. (1987). Factor analysis and AIC. Psychometrika, 52(3), 317-332. doi:10.1007/bf02294359
2. Bagozzi, R. P., & Edwards, J. R. (1998). A General Approach for Representing Constructs in Organizational Research. Organizational Research Methods, 1(1), 45-87. doi:10.1177/109442819800100104
3. Barak, G. (2008). Toward an Integrative Study of International and State-Corporate Criminality: A Reciprocal Approach to Gross Human Rights Violations. In R. Haveman and A. Smeulers (eds.), Supranational Criminology: Towards Criminology of International Crime. Antwerp, The Netherlands: Intersentia Press.
4. Bellair, P. E., Roscigno, V. J., & Mcnulty, T. L. (2003). Linking Local Labor Market Opportunity To Violent Adolescent Delinquency. Journal of Research in Crime and Delinquency, 40(1), 6-33. doi:10.1177/0022427802239252
5. Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88(3), 588-606. doi:10.1037/0033-2909.88.3.588
6. Bentler, P. M. (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107, 238–246.
7. Bernard, T. J., & Snipes, J. B. (1996). Theoretical integration in criminology. Crime and Justice: An Annual Review of Research, 20, 301-348.
8. Bernard, T. J., Vold, G. B., Snipes, J. B., & Gerould, A. L. (2010). Vold's theoretical criminology. New York: Oxford University Press.
9. Bik, & O.P.G. (2010). The behavior of assurance professionals: a cross-cultural perspective. Eburon.
10. Bollen, K. A. (1998). Testing structural equation models. Newbury Park, Calif. [u.a.: Sage.
11. Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika, 52(3), 345-370. doi:10.1007/bf02294361
12. Browne, M., & Cudeck, R. (1989). Single Sample Cross-Validation Indices for Covariance Structures. Multivariate Behavioral Research, 24(4), 445-455. doi:10.1207/s15327906mbr2404_4
13. Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. Psychological Bulletin, 105(3), 456-466. doi:10.1037/0033-2909.105.3.456
14. Carmines, E.G., & McIver, J.P. (1981). Analyzing Models with Unobserved Variables. In G.W. Bohrnstedt & E.F. Borgatta (Eds.), Social Measurement: Current Issues, Beverly Hills, CA: Sage.
15. Cain, M. (2000). Orientalism, Occidentalism and the Sociology of Crime. British Journal of Criminology, 40(2), 239-260. doi:10.1093/bjc/40.2.239
16. Cheung, G. W., & Rensvold, R. B. (2000). Assessing Extreme and Acquiescence Response Sets in Cross-Cultural Research Using Structural Equations Modeling. Journal of Cross-Cultural Psychology, 31(2), 187-212. doi:10.1177/0022022100031002003
17. Cole, D. A., & Maxwell, S. E. (1985). Multitrait-Multimethod Comparisons Across Populations: A Confirmatory Factor Analytic Approach. Multivariate Behavioral Research, 20(4), 389-417. doi:10.1207/s15327906mbr2004_3
18. Farrington, D. P. (2000). Explaining and preventing crime: the globalization of knowledge –the American society of criminology 1999 presidential address. Criminology, 38(1), 1-24. doi:10.1111/j.1745-9125.2000.tb00881.x
19. Fraser, A. (2013). Ethnography at the periphery: Redrawing the borders of criminology's world-map. Theoretical Criminology, 17(2), 251-260. doi:10.1177/1362480612472786
20. Garson, G. D. (2008). Structural equation modeling. In Statnotes: Topics in multivariate analyses. Retrieved from http://faculty.chass.nesu.edu/garson/PA765/garson.htm
21. Gregorich, S. E. (2006). Do Self-Report Instruments Allow Meaningful Comparisons Across Diverse Population Groups? Medical Care, 44(Suppl 3), S78-S94. doi:10.1097/01.mlr.0000245454.12228.8f
22. Hantrais, L., & Mangen, S. P. (1996). Cross-national research methods in the social sciences. London: Pinter Publ.
23. Hirschi, T. (1969). Causes of delinquency. Berkeley: University of California Press.
24. Hirschi, T. (1989). Exploring Alternatives to Integrated Theory. In Theoretical Integration in the Study of Deviance and Crime: Problems and Perspectives (p. 37–49). State University of New York Press.
25. House, R. J., & Global Leadership and Organizational Behavior Effectiveness Research Program. (2004). Culture, leadership, and organizations: The GLOBE study of 62 societies. Thousand Oaks, CA: Sage Publications.
26. Horn, J. L., & Mcardle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. Experimental Aging Research, 18(3), 117-144. doi:10.1080/03610739208253916
27. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1-55. doi:10.1080/10705519909540118
28. Hui, C. H., & Triandis, H. C. (1985). Measurement in Cross-Cultural Psychology: A Review and Comparison of Strategies. Journal of Cross-Cultural Psychology, 16(2), 131-152. doi:10.1177/0022002185016002001
29. Jöreskog, K. G., Sörbom, D., & SPSS Inc. (1981). LISREL 8 user's reference guide. Chicago, IL: Scientific Software International.
30. Kankaras, M., & Moors, G. (2012). Cross-National and Cross-Ethnic Differences in Attitudes: A Case of Luxembourg. Cross-Cultural Research, 46(3), 224-254. doi:10.1177/1069397112440945
31. Kaplan, D. (2000). Structural equation modeling: Foundations and extensions. Thousand Oaks, CA: Sage Publications.
32. Karstedt, S. (2001). Comparing Justice and Crime across Cultures. The SAGE Handbook of Criminological Research Methods, 373-390. doi:10.4135/9781446268285.n25
33. Klein, M. W. (1989). Cross-National Research in Self-Reported Crime and Delinquency. Dordrecht: Springer Netherlands.
34. Leighton, D., Harding, J. S., Macklin, D. B., Macmillan, A. M., & Leighton, A. H. (1963). The character of danger: Psychiatric symptoms in selected communities. (The Stirling County study of psychiatric disorder & sociocultural environment, 3.) New York: Basic Books.
35. Leung, K., Bond, M. H., De Carrasquel, S. R., Munoz, C., Hernandez, M., Murakami, F., … Singelis, T. M. (2002). Social Axioms: The Search for Universal Dimensions of General Beliefs about How the World Functions. Journal of Cross-Cultural Psychology, 33(3), 286-302. doi:10.1177/0022022102033003005
36. MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1(2), 130-149. doi:10.1037/1082-989x.1.2.130
37. Marshall, I. H., & Marshall, C. E. (1983). Toward a Refinement of Purpose in Comparative Criminological Research: Research Site Selection in Focus. International Journal of Comparative and Applied Criminal Justice, 7(1-2), 89-97. doi:10.1080/01924036.1983.9688765
38. Meade, A. W., & Bauer, D. J. (2007). Power and Precision in Confirmatory Factor Analytic Tests of Measurement Invariance. Structural Equation Modeling: A Multidisciplinary Journal, 14(4), 611-635. doi:10.1080/10705510701575461
39. Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. Psychometrika, 58(4), 525-543. doi:10.1007/bf02294825
40. Messner, M. (2011). Gender Ideologies, Youth Sports, and the Production of Soft Essentialism. Sociology of Sport Journal, 28(2), 151-170. doi:10.1123/ssj.28.2.151
41. Messner, S. F., Krohn, M. D., & Liska, A. E. (1989). Theoretical integration in the study of deviance and crime: Problems and prospects. Albany: State University of New York Press.
42. Milfont, T.L., & Fischer. (2010). Testing measurement invariance across groups: applications in cross-cultural research. Universidad de San Buenaventura (USB), Medellín.
43. Millsap, R. E., & Kwok, O. (2004). Evaluating the Impact of Partial Factorial Invariance on Selection in Two Populations. Psychological Methods, 9(1), 93-115. doi:10.1037/1082-989x.9.1.93
44. Muftic, L. R. (2009). Macro-micro theoretical integration: an unexplored theoretical frontier. Journal of Theoretical and Philosophical Criminology, 1(2), 33-71.
45. Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. Psychological Bulletin, 105(3), 430-445. doi:10.1037/0033-2909.105.3.430
46. Nelken, D. (2009). Comparative Criminal Justice: Beyond Ethnocentrism and Relativism. European Journal of Criminology, 6(4), 291-311. doi:10.1177/1477370809104684
47. Paternoster, R., & Bachman, R. (2001). Explaining criminals and crime: Essays in contemporary criminological theory. Los Angeles, CA: Roxbury Pub. Co.
48. Pratt, T. C., & Godsey, T. W. (2002). Social support and homicide: a cross-national test of an emerging criminological theory. Journal of Criminal Justice, 30(6), 589-601. doi:10.1016/s0047-2352(02)00192-7
49. Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. Psychological Bulletin, 114(3), 552-566. doi:10.1037/0033-2909.114.3.552
50. Ren, L., Zhao, J. S., He, N. P., Marshall, I. H., Zhang, H., Zhao, R., & Jin, C. (2015). Testing For Measurement Invariance of Attachment Across Chinese and American Adolescent Samples. International Journal of Offender Therapy and Comparative Criminology, 60(8), 964-991. doi:10.1177/0306624x14566602
51. Rensvold, R. B., & Cheung, G. W. (1998). Testing Measurement Models for Factorial Invariance: A Systematic Approach. Educational and Psychological Measurement, 58(6), 1017-1034. doi:10.1177/0013164498058006010
52. Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. Human Resource Management Review, 18(4), 210-222. doi:10.1016/j.hrmr.2008.03.003
53. Sekaran, U. (1983). Methodological and Theoretical Issues and Advancements in Cross-Cultural Research. Journal of International Business Studies, 14(2), 61-73. doi:10.1057/palgrave.jibs.8490519
54. Sjoberg, G. (1955). The Comparative Method in the Social Sciences. Philosophy of

Science, 22(2), 106-117. doi:10.1086/287408
55. Smelser, N. J. (1976). Comparative methods in the social sciences. Englewood Cliffs, NJ: Prentice-Hall.
56. Steenkamp, J., & Baumgartner, H. (1998). Assessing Measurement Invariance in Cross-National Consumer Research. J CONSUM RES, 25(1), 78-107. doi:10.1086/209528
57. Steiger, J. H., & Lind, J. C. (1980). Statistically based tests for the number of common factors. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA
58. Stodolsky, S. S., & Lesser, G. S. (1973). Learning patterns in the disadvantaged. Bethesda, MD: ERIC Reproduction Service.
59. Thompson, B. (2000). Canonical correlation analysis (pp. 285-316). In: Grimm, L. & Yarnold, P. (Eds.). Reading and understanding more multivariate statistics. Washington: APA.
60. Vallier, I., & Apter, D. E. (1971). Comparative methods in sociology: Essays on trends and applications. Berkeley: University of California Press.
61. Van de Vijver, & F.J.R. (1998). Towards a theory of bias and equivalence. (Zuma Nachrichten vol.3 (1998) p.41-65 [ISSN 0721-8516].
62. Van de Vijver, F. J. R., & Leung, K. (1997). Methods and data analysis for cross-cultural research. Newbury Park, CA: Sage.
63. Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. Organizational Research Methods, 3(1), 4-70. doi:10.1177/109442810031002
64. Warwick, D. P., & Osherson, S. (1973). Comparative research methods. Englewood Cliffs, NJ: Prentice-Hall.
65. Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing Reliability and Stability in Panel Models. Sociological Methodology, 8, 84. doi:10.2307/270754
66. Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. The science of prevention: Methodological advances from alcohol and substance abuse research, 281-324. doi:10.1037/10222-009
67. Yoon, M., & Millsap, R. E. (2007). Detecting Violations of Factorial Invariance Using Data-Based Specification Searches: A Monte Carlo Study. Structural Equation Modeling: A Multidisciplinary Journal, 14(3), 435-463. doi:10.1080/10705510701301677
68. Øyen, E. (1990). Comparative methodology: Theory and practice in international social research. London: Sage.