Perceptions of Developmental Hierarchies in Taiwan –

Conceptual, Substantive, and Methodological Insights

Arland Thornton
University of Michigan

Li-shou Yang
University of Michigan

Contact author: Arland Thornton, arlandt@umich.edu, 734/615-3542, Ann Arbor, Michigan 48106

Abstract

Motivated by a growing awareness of the penetration of world culture into the daily lives of ordinary people, this paper analyzes Taiwanese college students' perceptions of developmental hierarchies, a key element of models of modernization. We investigate the extent to which Taiwanese students hold hierarchical views of the world, whether these views match the views of the United Nations, the stability of these views across time, and the reliability of measurement. Data for this paper come from the survey of "Political Values and Attitudes among University Students in Taiwan", a panel study conducted by the Election Study Center in Taipei, Taiwan. Our results from this panel study conducted in 2006, 2007, and 2008 show that Taiwanese students have worldviews that include developmental hierarchies that are very similar to the country development ratings of the United Nations. We show that these perceptions of developmental hierarchies can be measured reliably at both the individual and aggregate levels and are stable across the survey years.

Introduction

Over the past decade there has been an expansion of interest in understanding the ways in which the interrelated ideas of development and modernization have been globalized and become important ideational factors in the lives of ordinary people around the world. This paper is designed to contribute to our understanding of this issue by focusing on a key element of models of development--the idea that the countries of the world form a developmental hierarchy. Our focus is on a series of interrelated substantive and methodological issues: whether people embrace a world view in which various countries are encapsulated into a hierarchy according to their levels of development; whether these views match the views of the United Nations; the stability of these views across time; and the measurement properties of questions designed to measure people's perceptions of developmental hierarchies.

Viewing the world as containing a developmental hierarchy is related to the ideas of development and modernization that have been important for scholars and policy makers in the western world for hundreds of years (Harris 1968, Mandelbaum 1971, Nisbet 1975/1969, Sanderson 1990, Thornton 2001, Thornton 2005). Development and modernity have also become important elements in a world culture that has been disseminated throughout much of the world through numerous mechanisms. This world culture has been shown to influence many aspects of societies, including the nature of the state, laws, public policies and programs, school organization and curriculum, and nongovernmental organizations (Baker and LeTendre 2005, Barrett and Frank 1999, Benavot, et al. 1991, Boyle 2002, Chabbott 2003, Drori and Krücken 2009, Frank, et al. 2010, Meyer, et al. 1997, Tsutsui and Wotipka 2004). The globalization of the ideas of development and modernization has been extended in many places to ordinary people in everyday life.

The documentation of this global dissemination of the ideas of development and modernity is important because such ideas provide schemas or frameworks for understanding and dealing with the world that are very different from the indigenous cultures or schemas that have existed in local societies for centuries. As a consequence, the dissemination of developmental models can be a powerful force for political, familial, demographic, and religious change. Such dissemination can also result in tension and conflict as the new ideas come into contact with powerful, previously-existing, and contrary schemas.

A recent but growing body of survey research has documented that the ideas of development and developmental hierarchies are widespread in many parts of the world (Binstock and Thornton 2007, Melegh, et al. 2012, Thornton, et al. 2012, Xie, et al. 2012). However, that research has only begun to answer the many conceptual, substantive, and methodological questions associated with beliefs about development and developmental hierarchies. Although this research has demonstrated that such ideas are often closely related to the conceptual and substantive views of such international organizations as the UN and are associated with such basic demographic factors as education and gender, it has not investigated the wide range of forces that might influence people's beliefs in developmental hierarchies. It has also not investigated the influence of such ideas on subsequent behavior, such as voting, migration, education, and family life. The research community has also not yet done the research to understand whether the distributions of reported beliefs about development are measured reliably at either the individual or aggregate levels. We also do not know the extent to which measures of developmental beliefs are stable at the individual or aggregate levels. In addition, there is no information about individual response styles in the ways that individuals answer survey questions about development, and, if there are, whether these response styles substantially bias data. This lack of information means that we do not know if the survey measures used to date provide reliable and stable measures that can be used to study the causes and consequences of developmental beliefs at the individual or aggregate levels. These limitations exist because this research is relatively new and because the survey research reported to date has been limited to one-time cross-sectional data collections.

The goal of this paper is to answer 15 interrelated substantive, conceptual, and methodological questions about people's perceptions on developmental hierarchies of countries. We divide these questions into four subgroups. Our first subgroup of guiding questions focuses on developmental awareness and the aggregate distributions of development ratings:

1. Are people able to rate countries on development?
2. Do people differentiate countries on their levels of development?
3. Are the average ratings of development for countries stable over time?
4. Are the detailed distributions of ratings for each country stable over time?

The next set of questions move us from the aggregate level to the individual level and address the reliability and stability of individual respondent ratings.

5. How reliable are individual raw ratings of country development?
6. How stable are individual raw ratings of country development?
7. How do reliabilities of individual raw ratings for country development compare to reliabilities of other survey measures?
8. How do stabilities of individual raw ratings of country development compare to stabilities of other survey measures?

The third set delves into the issue of individual response styles in answering questions about development.

9. Do individuals have different styles in the ways they answer country development questions?
10. How reliable and stable are individual rating styles?
11. How reliable and stable are the individual development ratings when adjusted for response styles?

We are also interested in the issue of how the development hierarchies held by ordinary people match or are different than the development hierarchies held by elite international organizations, such as the United Nations. This is important because a similarity of elite hierarchies with those of ordinary people would indicate that the elite models have permeated to the grassroots. This motivated us to compare our respondent development ratings with the UN human development ratings (HDI). More specifically, our last four research questions focus on the correspondence between respondent rating scores and UN HDI scores.

12. Do average individual raw development scores match the UN development scores?
13. Do average individual adjusted scores match the adjusted UN scores?
14. How do individual scores match the UN scores?
15. How reliable and stable are the correlations of individual ratings with the UN scores?

To our knowledge, our research is the first to examine this broad set of interrelated substantive, conceptual, and measurement issues, and we do so using a three-wave panel study of college students in Taiwan. One component of this Taiwanese panel study was designed to investigate the extent to which students understand and use the concept of development. Using the methodology created in earlier studies, we asked students to rate the development of several countries on an eleven point scale (Binstock and Thornton 2007, Melegh, et al. 2012, Thornton, et al. 2012, Xie, et al. 2012). The same questions were asked at each wave, permitting us to investigate important issues of reliability, stability, and individual response styles.

Our goal is not to endorse developmental models but to assess how people view development, whether those views match the views of the UN, whether those views can be measured reliably, and whether those views are persistent over time. Our interest in studying people's views of development and developmental hierarchies does not come because we believe that the developmental model is in some way true and useful for scholarly research and public policy. We recognize that there have been sustained critiques of modernization theory as an organizing framework for studying human life and for fostering public programs (Baker 1998, Boas 1940, Bock 1956, Böröcz 2000, Chakrabarty 2000, Comaroff and Comaroff 1992, Hodgen 1964, Jennings 1975, Mandelbaum 1971, Nisbet 1975/1969, Szreter 1993, Tilly 1984, Wallerstein 1991).

In the following, we present the conceptual and theoretical motivation of our study. We describe the ideas of development, discuss their dissemination to ordinary people around the world, and explain how they can be important in understanding individual decision making concerning several important behaviors. We then describe our three-wave panel study of Taiwanese students, describe our survey questions, and explain our methods for studying reliability, stability, and individual response styles. We then present our results and end with conclusions.

Conceptual Background

For centuries, the view of the world as a hierarchical system characterized by countries with different levels of development has been a common perspective in social science and political discourse. This viewpoint primarily stems from modernization or development theory. Early scholars generated a theory about societal development in which countries progress along the same trajectory but at different speeds (Harris 1968, Mandelbaum 1971, Nisbet 1975/1969, Sanderson 1990, Thornton 2001, Thornton 2005). Countries around the world could thus be conceived as situated in a developmental hierarchy in which the position of each country was determined by its progress along the developmental path. Therefore, at any time, one could perceive various countries at different levels of development. The countries of northwest Europe and their overseas populations were generally perceived to be at the apex of this developmental hierarchy. In addition, the model assumed that development was a universal outcome that was attainable for any society by following the path of those seen as further advanced. Many developmental theorists also considered development as good, desirable, and a worthy goal to pursue.

This developmental model has dominated much social science research and political discourse for centuries. It has also become an important element of world culture and has been spread globally. One example of the international scope of this developmental model is the practice of many national and international organizations such as the United Nations, the World Bank, and the International Monetary Fund to divide countries into groups based on their development levels (International Monetary Fund 2009, United Nations Development Programme 2009, World Bank 2010).

This development model provides a comprehensive framework to explain past history, categorize present circumstances, and prescribe the future direction for societies. This theoretical thinking has built the foundation for numerous development projects initiated and implemented by social organizations, political entities, and international agencies. Its influence can be exemplified in the policies of colonial governments in the 18th and 19th centuries, in the programs of the Meiji Restoration in Japan (1868), in the renewal movement in China in the late 19th and early 20th centuries, in the establishment of the Republic of Turkey in the early 20th

century, and in revolutions and reforms in many other places. In addition, many in the feminist movement, family planning advocates, and other social movements have relied heavily on developmental models. The influence of the model has been especially profound after World War II, owing to the increasing number of international development organizations, the expansion of world society, and the openness of globalized networks.

The idea of developmental hierarchies has also been important in the relations between nations. Western Europeans have long viewed their region as more developed than the rest of Europe, and this perception was important in the way Europe was divided after World War II between "east" and "west" (Böröcz 2006, Melegh 2006, Todorova 1997, Wolff 1994). Perceptions of this developmental hierarchy have also influenced how Eastern European countries were evaluated for possible admission into the European Union (Böröcz 2000). The model has also greatly affected the relationships between Europeans and the non-European countries that Europeans consider to be less developed.

There is a growing body of evidence indicating that the ideas of development have not been limited to academics and public policy elites but have been globalized to ordinary people in everyday life. There have been numerous mechanisms for the spread of these ideas, including education, mass media, colonialism, social movements, foreign aid programs, and government and nongovernment efforts (Thornton 2005).

We know from historical and ethnographic studies that the basic ideas of development are understood by at least some ordinary people in China, Egypt, India, Nepal, New Guinea, and parts of Sub-Saharan Africa (Abu-Lughod 1998, Ahearn 2001, Amin 1989, Caldwell, et al. 1988, Dahl and Rabo 1992, Ferguson 1999, Guneratne 2001, Justice 1986, Liechty 2003, Osella and Osella 2006, Pigg 1996, Wang 1999). For example, Ferguson (1999, p. 14) has written that "the narrative of modernization ....gives form to an understanding of the world, providing a set of categories and premises that continue to shape people's experiences and interpretations of their lives." He reported that the conceptions of urban African workers were "not simply *compatible* with the modernist narratives of social science; they were a local version of them. Modernization theory had become a local tongue…" (p. 84). Similarly, Osella and Osella (2006, pp. 570-571), two anthropologists who study the Arabian Peninsula and South Asia, have concluded that "concepts, ideals, and practices of something called 'modernity'…exist and are

continually appealed to in people's economic endeavours, political projects, and identity crafting… 'Modernity', as an historically and ethnographically specific body of ideals and practices, will remain central to our understanding of contemporary societies."

Evidence is also accumulating from survey research showing that ordinary people in everyday life understand and use the basic ideas of modernization theory and hierarchical development models. Thornton and his colleagues (2012a) used data from fifteen surveys in thirteen diverse countries to show that ordinary people in each setting understand developmental hierarchies and can use this understanding to rate a series of countries on development. In addition, large percentages of people in these thirteen countries rate countries on development similarly to the development ratings of the UN HDI, suggesting that development is not only widely understood, but understood in ways that are similar to the perspectives of international elites (Binstock and Thornton 2007, Melegh, et al. 2012, Thornton, et al. 2012, Xie, et al. 2012)

This body of research has also shown that while views of development and developmental hierarchies are widespread among individuals in many countries, there is also considerable variability in the ways people perceive development (Thornton, et al. 2012). This is demonstrated in the fact that significant proportions of people either do not subscribe to developmental hierarchies or define them differently than does the United Nations and other international elites. Having different views from the UN is not surprising because development is a social construct that does not exist in an objective sense beyond the discourse itself (Sumner and Tribe 2008). This can be exemplified by the fact that, from time to time, the concept of development has been redefined and its indicators have been modified (Haq 1995, Miles 1985, United Nations Development Programme 2010). This makes it important to understand how individual people use the language and imagery of development to characterize their own and other countries. The issue then becomes how these images are produced and what their social, political, and economic effects are, rather than on whether or not these images are "true" in some objective sense. The ways in which individual views of development and the international developmental hierarchies do and do not match those of the UN can provide insights into people's social constructions.

It is important to recognize that the acceptance, rejection, or modification of developmental models can be important influences on the decisions and behavior of ordinary people in a wide array of areas, including politics, gender relations, health, family, and demography. We expect that individuals who accept developmental models, including their international hierarchies, will behave differently from individuals who reject or substantially modify the models. This potential for influencing behavior does not depend on whether the elements of the model, themselves, are true or false, good or bad, but whether they are believed. Our issue, therefore, is not the validity or goodness of the developmental models, but whether they are accepted, rejected, or modified by individual people.

Understanding the influence of these developmental models in the lives of ordinary people requires that the research community have a battery of reliable indicators. Such survey indicators must be able to distinguish reliably between people with different views and understandings of the models. They must also be able to take into account the basic response styles that people use to answer survey questions about development.

As we discussed in the introduction, our aim in this paper is to examine in detail an interrelated set of fifteen substantive, conceptual, and measurement issues concerning people's views of development and developmental hierarchies. We do so by analyzing data from a three-wave panel study of college students in Taiwan to determine how the students differentiate countries on their levels of development. We examine how these views of students match those of the UN, whether the perceived differentiations of country development can be reliably ascertained through survey questions, whether there are personal response patterns in answering these questions, and whether people's views are stable across time.

We recognize that our data come from a specialized sample of respondents from a single country, which limits the generalizability of our results. Nevertheless, we believe that our analysis is valuable as a first evaluation of these issues. It provides both initial estimates of these issues and a foundation and motivation for additional research using more general populations.

Data and Measurement

The panel data for this analysis come from the Study of Political Values and Attitudes among University Students in Taiwan, conducted by the Election Study Center in Cheng-Chi University, Taipei, Taiwan. Based on equal probability sampling, the baseline for the panel study was conducted in October 2004 and interviewed 80 percent of freshmen students (sample size=1,368) of the National Cheng-Chi University in Taiwan. These students were re-interviewed annually in April of the following four years (2005, 2006, 2007, and 2008). Through self-administered questionnaires, this panel study collected respondents' ratings for the development levels of several countries in the 2006, 2007, and 2008 follow-ups.

Because our study focuses on perceptions over three waves of repeated measurement, we limit our analyses to those respondents who consistently remained in the study through the three waves of follow-ups. As in all panel studies, cumulative non-responses over waves led to attrition in the panel data used in this paper[1]. Among the original sample, 998 respondents (about 73% of the baseline sample) were re-interviewed in the 2006 follow-up and 762 (about 57% of the baseline sample) were re-interviewed through 2008.

The perceptions of country development in this panel study were measured on a zero-to-ten scale, with ten representing the highest development and zero representing the lowest. The countries rated consists of the following ten countries listed according to their order in the questionnaire: Japan, Nigeria, India, USA, China, Zimbabwe, Sweden, Brazil, Pakistan, and Taiwan. In order to document the developmental hierarchy in the minds of ordinary people, neither the meaning of development nor a description of the countries was provided to respondents in the surveys. This let respondents freely conceptualize the development levels of countries according to any ideas and indicators within their awareness. Respondents were instructed to use their best estimate to rate the development of a country when they believed that they lacked information.

To gauge the possible impact of sample attrition on the analysis, we compared the average ratings reported by all respondents in the 2006 follow-up and those reported by

---

[1]Attrition is inevitable in any panel study (Alwin 2007). It is not unusual for a study to carry out panel analysis based on less than one-half of the original sample (see Coenders, et al. 1999).

respondents remaining throughout 2007 and 2008. The first two columns in Table 1 are the average ratings measured in 2006 for the 2006 sample, in contrast to the ratings of the restricted 3-wave panel sample from 2006 to 2008. The results show almost identical average ratings reported by the full and restricted samples. Therefore, even though the panel study is unable to retain all respondents over three waves, we can confidently use the panel sample for the analysis without sample attrition being a major problem.

The development scores given by respondents for several countries based on the 11-point scale are useful in reflecting people's evaluations of the developmental levels of a set of countries at a particular time point. Nevertheless, it has been observed by earlier survey researchers that respondents differ in their tendency of choosing response categories in a scale (Cronbach 1950, Cunningham, et al. 1977). When applying a fixed-point scale to rate developmental levels, some people may tend to rate all countries with high scores, while others may tend to give all countries low scores. In addition, some people may tend to avoid the response categories in the two extremes of a scale, while others use the entire range of the scale. The observed differences in ratings provided by different individuals could thus be partially affected by the inter-personal differences in rating styles. In addition to differences in rating styles between individuals, there is possible variation in the rating styles of individuals over time. The inter-personal differences or individuals' over-time changes in rating style are independent from the content of questions and generate noise in measurement. In other words, response styles are likely to generate confounding effects in the inter-personal comparison of developmental scores. Also, the comparison of ratings provided by an individual at different time points could be confound by a person's change in rating styles across waves.

For these reasons, we employ two versions of a respondent's rating of a specific country in the developmental hierarchy. The first is the raw score that a particular respondent gave for that specific country and is based on the original eleven-point scale in the survey. The second is an adjusted or standardized score that takes into account whether an individual tends to rate all countries relatively high or low and whether an individual tends to rate countries in a narrow or wide range. The standardized scores of countries are calculated separately for each individual taking into account that individual's mean and standard deviation for his/her ten

ratings. That is, to calculate the standardized score for a specific country for a particular individual, we subtracted that particular individual's mean score for all countries from that individual's score for the specific country and divided the result by the individual's standard deviation across all countries.

Through this transformation of raw scores into standardized or adjusted scores, regardless of individual differences in response styles, every country rated can be conceived as situated in an individual's ten-country hierarchical system where the average developmental score for these ten countries is zero and the standard deviation is one. In other words, the standardization of developmental hierarchies generates a series of converted scores which measure developmental levels of countries within a hierarchy with a uniform mean and standard deviation across individuals.

Analysis Methods

We examined whether individuals rate the development of various countries in a hierarchical way and whether the developmental hierarchies expressed by individuals are stable over time. While the cross-time correlations of repeated measures for people's perceptions of developmental hierarchies are indicators of the consistency of the perceived developmental hierarchy, they reflect both over-time stability and reliability of measurement. Therefore, we use a procedure to disentangle the reliability of measurement and the stability of the variable.

We apply a linear structural equation method incorporating latent unobserved variables to estimate reliability and stability for panel data, as proposed by Heise (1969) and Alwin (2007, 1989). This model, summarized in Figure 1, simultaneously estimates the measurement reliability and the cross-time stability of an underlying construct. The model can be partitioned into two parts, each presented by a set of equations. The first part is the measurement model with three equations linking indicators with the underlying or true variables. We posit that each individual has a single underling true score for each country's development (η), which has a single indicator (y) that is observed in three occasions, the 2006, 2007, and 2008 interviews. Each rating for a country's development (y_i) for each of the three years is assumed to be linked

to the underlying construct in the same year (η_i) through λ_i, as presented in Equations 1, 2, and 3. The lambda coefficient (λ_i) reflects the effect of η_I on y_i. The observed indicator is assumed to have two components, the true score (η_i) and the error of measurement (ε_i). We apply a standard assumption in classic test theory that the error, ε_i, is randomly distributed—that the ε_i, are uncorrelated with each other and with η_i.

Change or stability in the perceptions of development of various countries is also of interest in this analysis. There are several factors which could contribute to the true change of the responses. One of the factors is the increase of information over time. Students could differentially obtain information and knowledge about countries through college educations, newspapers, the internet, and other sources. Therefore, the second part of the model incorporates Equations 4 and 5 that take into account the possible changes of underlying variables over time. The model assumes a Markovian process (lag-1) which in essence represents a simple ordering scheme. In this process, the occurrence of the underlying variable at time t is assumed to be dependent only on the underlying variable at time t-1 with an error of prediction (ζ). The ζ's are assumed to be independent.

$$y_1 = \lambda_1 \eta_1 + \varepsilon_1 \qquad (1)$$
$$y_2 = \lambda_2 \eta_2 + \varepsilon_2 \qquad (2)$$
$$y_3 = \lambda_3 \eta_3 + \varepsilon_3 \qquad (3)$$
$$\eta_2 = \beta_{21} \eta_1 + \zeta_2 \qquad (4)$$
$$\eta_3 = \beta_{32} \eta_2 + \zeta_3 \qquad (5)$$

This model estimates both reliability of measurements and stability of true scores over time. The concept of reliability represents the extent to which a response given to a survey question is "true" variation rather than variation produced by measurement error. Therefore, it is defined as the ratio of true-score variance to observed score variance (σ^2_τ/σ^2_y) and operationalized as the proportion of the variance of the underlying variable (η) to the variance of the observed variable (y): reliability of $Yi = V(\eta_i)/(V(\eta_i) + (V(\varepsilon_i))$. The reliability of a measure is also equal to the square of that measure's estimated lambda coefficient.

Figure 1. Single-Indictor Reliability and Stability Model
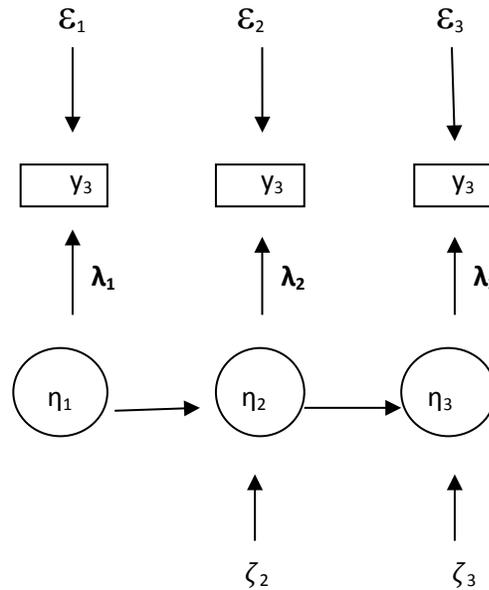
Stability refers to the continuity of the underlying construct. Since each equation has only one predictor variable, the βs are also equal to the correlations between the two observations of the same construct at two points in time. The βs are referred to as stability coefficients (Alwin 1989).

With only three observations, it is impossible to estimate all the parameters of these equations without making some assumptions. Following the approach of Heise (1969) and Alwin (1989), we assume equal reliability over time. This is a useful and realistic assumption when the observed variances do not vary with time. Detailed discussions of the application of constant reliability across waves and the Markovian restriction in the assessment of the reliability of a single indicator and the stability of underlying constructs can be found in Alwin (1994, 2007, 1989) and Jöreskog (1974, 1970).

Results

We begin our presentation of results with our first four guiding questions. These questions focus on development awareness and the average ratings and detailed distributions of development ratings of ten countries.

1. Are people able to rate countries on development?

The initial step of our analysis examines whether survey respondents are able to rate countries on development. Our data answer this question positively; a majority of respondents were able to provide ratings for all ten countries in each of the three waves. Over three waves, there were only 14 respondents who were unable to provide a complete set of ratings for the ten countries. Each of these 14 people was unable to give Zimbabwe a score in at least one interview. Two of these 14 respondents were unable to rate development for multiple countries. Thus, the vast majority were able to provide ratings for ten countries in all three waves.

Five respondents in 2006, one in 2007, and two in 2008 rated all countries with the same scores. In total, seven respondents spreading across three waves provided invariant developmental ratings for the ten countries in one of the three waves. That means, in each wave, almost all respondents differentiated between countries in levels of development, rating some countries high and some countries low. Thus, there is an imagined hierarchy in assessments of development for countries around the world, although we have no information about the criteria Taiwanese students used to arrange the countries from high to low development.

In the following analysis, we exclude the cases with missing data and with invariant ratings. We also exclude three cases that applied the scale in the opposite direction in one of the three waves. Two cases with high kurtosis indices causing difficulty in the estimation of structural equation models are also excluded. We also excluded one person who provided ratings for the ten countries as 1, 2, 3, 4, …. The total number of respondents included in the following analysis is 734.

2. Do people differentiate countries on their levels of development?

We next ask whether respondent ratings of countries form a hierarchy of development. In Table 1 (last three columns), we show for each wave the average ratings of each country rated by the 734 respondents. These data show that in each wave the average ratings form a collective perception of a developmental hierarchy.

On average, Taiwanese college students consistently gave the United States the highest developmental scores among the ten countries. In the first two waves, Japan was rated by these

students as the second most developed country, and Sweden was rated as the third most developed. In the last wave, the average rating for Sweden was nearly identical to that of Japan. Taiwan is in the fourth rank in the developmental hierarchy in all three waves. China is consistently ranked in the fifth place, followed by Brazil, then India, and then Pakistan.   The two African countries, Nigeria and Zimbabwe, received the lowest and the second lowest average scores.

3. Are the average ratings of development for countries stable over time?

As implied in the above discussion, the aggregate ratings of the ten countries are remarkably stable across survey waves. Most of the average ratings fluctuate only slightly over the three years. However, across the three-year period, the average ratings for India, Sweden, and Taiwan gradually increased. The increase for the ratings of Sweden is the most noticeable, being .38.

To summarize the remarkable aggregate stability of ratings, we present in the second panel of Table 1 the correlations among the distributions of average scores given by respondents in the three years. The average scores reported in 2006 and 2007 are almost perfectly correlated, as the Pearson correlation coefficient reaches .9997. The other cross-wave correlation coefficients are also above .99. These almost perfect correlations demonstrate that the collective perceptions of the developmental hierarchy are exceptionally consistent over time. Taiwanese college students have a world view with a developmental hierarchy, and that world view, at least in the aggregate, is very persistent across time.

4. Are the detailed distributions of ratings for each country consistent over time?

We now shift from an examination of average country ratings to an examination of the entire distribution of ratings for each country and ask whether the distribution of ratings for each country remains constant across waves. The data show that the distributions of ratings for each country's development are remarkably consistent across the three waves (not shown in tables). That is, the percentage of respondents giving a country any particular score is almost identical across waves.

The ratings for Japan, the United States, Taiwan, and Sweden are more tightly congregated around their mean scores than any other country's ratings (not shown in tables).

The ratings for Japan, the United States, and Sweden are all concentrated at seven or higher, with very few responses scattered below seven. In all three waves, more than 90 percent of ratings for Taiwan fall within the range of six to eight with close to 50 percent of the scores being seven. The ceiling effect—bounded by the highest developmental score at 10 -- causes the ratings for Japan, the United States, and Sweden to be less dispersed. Other than the ceiling effect, the commonly shared information about a country is also a plausible reason for the students to provide more congregated ratings for some countries. For example, the students in Taiwan have first-hand commonly shared information about the societal conditions where they live. Therefore, the ratings for Taiwan are intensively grouped around seven.

The ratings for India, China, and Brazil show relatively symmetrical distributions at the center of the 11-point scale but with different peak points (not shown in tables). Zimbabwe and Nigeria are rated with scores distributed toward the lower end of the developmental scale. About 60 percent of the ratings for Zimbabwe are spread between three and five. A similar percentage of ratings for Nigeria are concentrated at two or three.

5. *How reliable are individual raw ratings of country development?*

We now shift from the aggregate level to the individual level of analysis. Through the structural equation model presented in Figure 1, we estimated the reliabilities and stabilities of individual ratings based on the original 11-point scale of country development. The results are displayed in Table 2 and show that the reliabilities of the individual raw scores measuring the development of each of the ten countries are in the range of .65 (for Nigeria) to .43 (for Pakistan). The average reliability across the ten countries is .52. The data thus indicate that the raw scores for these countries are capturing substantial amounts of true variance in that 65% to 43% of the observed variance in the ratings represents true variance in the underlying variables. On the other hand, the data also reveal that a significant portion of the observed variance of the raw scores, 35% to 57%, can be attributed to noise or random error.

As we noted earlier, the estimated reliabilities are equal to the square of the lambda coefficients linking a variable's true score to its observed measure. This also means that the lambda coefficients are equal to the square root of the estimated reliabilities. Thus, the lambda coefficients range from .81 (for Nigeria) to .66 (for Pakistan), which represent substantial effects of true scores on observed scores.

6. *How stable are individual raw ratings of country development?*

Table 2 also displays the estimation indicating the extent to which the individual perceived developmental levels for each country remained stable between two adjacent interviews. The stability estimates for the raw scores are all above .8 for the two between-wave periods, with the average stability between 2006 and 2007 equaling .89 and the average stability between 2007 and 2008 equaling .90. This indicates a high degree of stability across these two one year periods.

7. *How do reliabilities of individual raw ratings for country development compare to reliabilities of other survey measures?*

In order to examine the comparative quality of data measuring developmental hierarchies and other survey measures, we estimated reliabilities and stabilities of nine variables measuring the social and political attitudes of the same panel respondents. These estimates, however, come from the baseline survey in 2004 and the 2005 and 2006 follow-ups rather than the 2006, 2007, and 2008 surveys that collected our development ratings. The comparisons we make are limited to attitudinal measures with the same eleven-point scale as the ratings for country development. We do so because the number of response categories could affect the reliability level (Alwin 2007). We estimated the reliabilities and stabilities of the following nine attitudinal measures: (1) homogeneity of ethnic relationship, (2) effectiveness of the electoral system, (3) satisfaction about democratic system, (4) optimistic attitude toward democratization, (5) preference between environmental preservation or economic development, (6) preference between social stability or social reform, (7) preference between independence or unification with China, (8) preference between traditional culture or foreign culture, and (9) attitude toward trading relations with Mainland China. The survey questions for these nine variables are presented in Appendix One.

As shown in the lower panel of Table 2, the reliabilities of these attitudinal variables ranged from .38 to .60, a range that is quite similar to the reliabilities of the scores measuring the developmental levels of countries. The lowest reliability for the developmental scores is for Pakistan which is .43, while the lowest reliability among the attitudinal measures is that for the preference of reform vs. social stability, which is .38. The highest reliability observed among the attitudinal variables is .60, for the optimistic attitude about democratization, which is lower than the highest estimation of reliability for developmental ratings, which is .65 for Zimbabwe.

The average reliability of the attitudinal variables is .51, while the average reliability for the developmental measurements is .52. Thus, in general, the developmental measures have very similar reliability to attitudinal measurements using an 11-point scale.

We also compared the reliabilities of the developmental measures with the reliabilities of variables of various domains from other panel surveys reported by Alwin (2007). Alwin applied the same approach that we used in this analysis to estimate the reliabilities of survey variables from six American panel studies. The variables in his study included many frequently used survey measures for belief, attitudes, values, and perceptions. The average reliability reported in his study for variables measuring beliefs is .48, for the variables measuring values, it is .56, for the attitudinal variables, it is .60, for the self-assessments, it is .50, and for self-perceptions, it is .50 (Alwin 2007, p 131). The average of the reliabilities for the variables from these five domains is .53. The mean estimates for the reliabilities of developmental ratings in our study is .52, which is very close to the average reliabilities in Alwin's study. The average reliability of our measurements for developmental hierarchies exceeds the average reliabilities of social survey measures in beliefs, self-assessments, and self-perceptions, but is lower than the average reliability of measures for values and attitudes reported by Alwin. This comparison, therefore, demonstrates that the reliability of our measures for developmental hierarchies are comparable with other frequently used ideational survey measures.

It is important to note that our data come from college students in Taiwan, while Alwin's data come from general populations in the United States. We do not know the extent to which country differences or the use of a college sample versus general samples might affect reliability estimates.

8. How do stabilities of individual raw ratings of country development compare to stabilities of other survey measures?

We now turn to the estimates of over-time stability of the attitudinal variables collected in the 2004 baseline and 2005 and 2006 waves of our survey. Because the time period between the 2004 baseline and 2005 follow-up was only six months, whereas all other inter-wave periods were twelve months, the stabilities from 2004 to 2005 are not directly comparable with the other stabilities estimated. Thus, we focus on the attitudinal stabilities from 2005 to 2006.

As shown in Table 2, the stabilities of the attitudinal items are somewhat lower than the stabilities of the development ratings. The least stable attitudinal measure between 2005 and 2006 is found for the optimistic view for democratization, while the least stable development score either between 2006 and 2007 or between 2007 and 2008 is for Zimbabwe, at .81 between 2006 and 2007. The average stabilities for the development measures are .89 for the period from 2006 to 2007, and .90 for the period from 2007 to 2008. The mean stabilities of the attitudinal variables are .86 during the 2005 - 2006 period. These stabilities indicate that the measures for country development in general are slightly more stable than the attitudinal variables.

9. Do individuals have different styles in the ways they answer country development questions?

Now we shift our focus to analyze the response styles shown by the set of development scores given by Individuals and the impact of response styles on the raw ratings based on the 11-point scale. We examined the different rating styles among individuals by calculating the minimum, maximum, mean, and standard deviation of the respondents' ratings of the ten countries. The results are presented by percentiles in Table 3.

The data provide evidence for the existence of different rating styles among respondents. We observe substantial inter-personal differences in both the average development scores and the degree of dispersion of the scores for the array of countries. In the 2006 data, the highest mean given by any respondent is 8.2 and the lowest mean is 4.1. While five percent of the respondents had average scores of the ten countries lower than 5, another five percent had average scores for the ten countries equal to or higher than 6.9. The distributions of mean scores for 2007 and 2008 are very similar to the data from 2006.

The standard deviation indicating dispersion of the ten-country ratings given by each respondent also varied among respondents. Some students arranged countries in a wider range of developmental hierarchy than others. The data from 2006 reveal that the standard deviations of individual ratings ranged from 1.06 to 4.22. The range from the fifth to the ninety-fifth percentile was from 1.63 to 3.37. Similar individual differences in dispersion existed in the other two waves.

10. *How reliable and stable are individual rating styles?*

Having demonstrated inter-person variation in rating styles, we now ask whether these rating styles are stable and measured reliably. We estimated the reliabilities and stabilities of the means and standard deviations obtained from the developmental ratings of the ten countries given by individual respondents.

A significant proportion of the observed differences of the individual mean scores reflect true differences in the mean scores of the ten countries given by respondents. The reliability of the mean ratings for the ten countries is .57 (not shown in tables), which is higher than the reliabilities for most of the other individual measurements we have estimated in this study. The reliability of the individual standard deviations is even higher and reaches .66 (not shown in tables). This reliability exceeds the estimated reliability of any other individual measurement in this study. Also the over-time stabilities of these two indicators range from .89 to .95 (not shown in tables), which are equal to or just a bit higher than the average stabilities of the developmental ratings shown in a previous section.

Therefore, a great proportion of observed variations in the rating styles indicated by both the mean and standard deviation of the set of ten ratings are true inter-personal differences which are independent of measurement errors. The stability estimates lead us to believe that individual rating styles are rather stable across survey waves. In other words, the differences in the ways in which respondents apply the 11-point scale in rating country development is very stable over one-year periods.

11. *How reliable and stable are the individual development ratings when adjusted for response styles?*

We have shown that different response styles exist in the ways individual respondents apply the 11-point scale to rate country development. The reliability and stability of these response styles affect the responses given in rating countries and probably affect the reliability and stability of the raw rating scores previously examined. This leads us to ask how reliable and stable the country development ratings are when the response style effects are removed.

As discussed earlier, to eliminate the possible influence of the differences in response patterns, we standardized each student's rating for each country based on the mean score and the standard deviation of the ten country ratings given by that particular student. Through the standardization process, the array of country ratings by each respondent is adjusted for that respondent's personal style in applying the 11-point scale. Using the same method used in other parts of the analysis, we estimated the reliability and cross-wave stability of the individual standardized ratings for country development. The results are presented in Table 4.

The estimates of reliabilities of the standardized scores ranged from .64 (for Zimbabwe) to .29 (for Pakistan). Except for the ratings for Taiwan and Pakistan, the reliabilities of the standardized ratings are all above .40. Except for Zimbabwe, the country standardized development scores are not as reliable as those measured by raw scores presented in Table 2. The average reliability of the standardized measures is .44 which is smaller than the average reliability of the raw scores, .52.

Comparing the cross-wave stability of individualized standardized ratings for country development, we find that the lowest cross-wave stability is .63 which is the stability for the rating of Japan in the period between 2006 and 2007. Combining the stabilities of two periods together, the standardized rating for Zimbabwe and Japan are the least stable among the ten country-measures, which is on average around .70. With only a few exceptions, the stabilities for the country development scores measured by standardized scores documented in Table 4 are lower than the stabilities of raw scores presented in Table 2.

The observation that the reliabilities and stabilities of the individual raw scores are higher than the reliabilities and stabilities of the standardized scores reflects the influence of respondent styles in rating countries on an 11-point scale. Individuals were different in their rating styles and this difference is relatively stable over time. That means there are habitual patterns of individuals in applying the scale to rate different counties. The habitual patterns have influence on the observable differences in the developmental ratings measured by the 11-point score. In other words, the estimated reliabilities and stabilities of raw scores measuring country development are inflated by the reliability and stability of individual rating styles.

In contrast to raw scores for the development ratings, the standardized scores take the individual variation in rating styles into account. The standardization process eliminates the influence of variation in rating styles. Therefore, the reliability and stability estimations for the standardized scores are not entangled with the reliability and stability of rating styles. The data show that after controlling the influence of rating style, the average reliability of measurement for developmental ratings reduces from .52 to .44, which is about a 15% reduction (.08/.52). We also find that after controlling rating styles, the average cross-wave stability of the ratings decreases from .89 to .84 between Wave 1 and Wave 2, and from .90 to .86 between Wave 2 and Wave 3 (reductions of 6 and 4 percent).

12. Do the average individual raw development scores match the UN development scores?

We next shift our examination from the reliability and stability of individual country development ratings and rating styles to an examination of how both aggregate and individual ratings compare with country development ratings made by elite organizations. Since 1990, in its annual Human Development Report the United Nations has provided numerical indices gauging the developmental levels of countries around the world. The UN Human Development Index (HDI) is formulated as a composite indicator, including measures of living standard, longevity, and education. In the following analysis we compared the survey reports of country development with the 2004 HDI to examine whether the developmental hierarchy defined by the experts of the UN fits with the views of ordinary people.

Note that the UN HDI scale is calibrated from zero to one, and in order to make the HDI scale intuitively more comparable to our respondent rating scale from zero to ten, we multiplied the HDI scores by ten. As we discuss more later, both the HDI scale and the respondent rating scale are arbitrary and different, meaning that the scores are only roughly comparable.

We first examine to what extent the students' average ratings for the ten countries are in agreement with the 2004 UN HDI. Because of the high stability of respondents' ratings for country development over time, in Table 5 we present only the comparison of the UN data with the data from our 2007 wave. The average scores given by Taiwanese students for the ten countries are consistently lower than the HDI scores of the same countries. As shown in the

table, the average student developmental scores for the ten countries are in the range from 2.88 to 9.20 in 2007, while the UN HDI scores for these countries ranged from 4.48 to 9.51. In addition, we consistently observe larger gaps between the highest and lowest scores for the students' responses than for the HDI scores. The lower means and higher ranges for the student reports indicate that their scales, on average, are calibrated differently than the one used by the UN, which is not surprising given that they originate in very different ways.

What is surprising is how close the average student reports are to the UN scores. In fact, the average scores reported by the students are correlated with the HDI scores at the remarkable level of .95. Thus, Taiwanese students, as a whole, have configured developmental hierarchies extremely similarly to the views of the UN experts.

13. Do the average individual student adjusted scores match the adjusted UN scores?

Because the UN HDI scores and our respondent raw score averages are calibrated on different scales, the raw score comparisons in Table 5 are only roughly comparable. For this reason, we also present in Table 5 the average standardized respondent scores for each country collected in the 2007 wave survey and UN HDI standardized scores. The standardized UN HDI scores are calculated based on the mean and standard deviation of the UN HDI scores of the ten countries. These results provide us with additional perspectives on the comparisons between the developmental hierarchy perceived by the Taiwanese students and the UN hierarchy.

Comparing the standardized scores obtained from UN HDI and from the respondents' answers, we found that some countries are, on average, rated higher by the students than by the UN HDI and that students rate other countries lower than does the UN. Higher average standardized estimates by students are particularly apparent for the United States and Zimbabwe. The standardized ratings for Taiwan, Brazil, and China are lower than the standardized UN HDI scores. The magnitudes of low estimation by the respondents are particularly noticeable for the ratings for Brazil and Taiwan which are both rated lower by .5 by the respondents than by the UN experts in the standardized ratings. The average standardized ratings given by the respondents for China are about .34 lower than the standardized UN score for China.

In a cross-country study of knowledge and beliefs about developmental hierarchies, Thornton and co-authors (2012a) found that respondents from most countries tend to rate the development of China highly. Our analysis, however, shows that even after standardization, the average scores for China provided by Taiwanese students are lower than the standardized UN score for China. In light of the past history and recent intensive exchanges and interactions between Taiwan and China, we believe that the circumstances in China are relatively well-known by most people in Taiwan. As the respondents were asked in self-administered questionnaires to rate the developmental levels of an array of countries including both Taiwan and China, they probably put the two in comparison and rated the two countries in correspondence to their understanding about the gap between the two places. Therefore, the low student ratings for Taiwan probably facilitated the correspondently lower rating for China. Examining the gap between the ratings for Taiwan and China given by the respondents and by the UN HDI, we find that the actual gap between these two countries is smaller from the survey reports than from the UN report. The average difference between Taiwan and China reported by the respondents is .61, while the gap shown in the UN report is .77.

14. *How do individual scores match the UN scores?*

We next examine the relation between the individual respondent's ratings for the ten countries with the UN HDI of these countries. We calculated the correlation between the ten ratings provided by an individual in each wave with the 2004 UN HDI of these countries. So, for each wave, we obtained 734 correlations, one correlation representing the congruence of that individual's report with the UN standard. The results are summarized by the percentile distribution of the Pearson correlations and presented in Table 6.

The data show that across all three waves, ratings provided by individual respondents are strongly correlated with the UN HDI. Only 5 percent of the respondents gave ratings for the ten countries that are correlated with the UN HDI scores at .73 or below. In all three waves, over 50% of respondents provided ratings that correlated with the UN HDI at .88 or higher. Thus, the high correspondence between average student ratings and UN ratings is not just the result of compensating errors but permeates the thinking of almost all Taiwanese students.

15. *How reliable and stable are the correlations of individual ratings with the UN scores?*

Finally, just as we estimated the reliabilities and stabilities of the various elements of the student country ratings, we estimated the reliabilities and stabilities of the correlations between the UN and individual student scores. The estimated reliability for the correlation between individual reports of the ten country developmental scores and the UN development scores is .36 (not shown in tables). A significant part of observable variance, about 64%, is due to the random error of measurement. The cross-wave stabilities for the correlation between the UN scores and the set of scores given by individual respondents are .85 or higher in the two between-wave periods (not shown in tables). These estimates suggest that the reliability and stability for the measure of congruence between the individual ratings and the UN HDI are relatively lower than the estimates for other variables. At the same time, they indicate substantial levels of both reliability and stability in this measure.

Summary and Conclusion

This paper began with the observation that modernization theory and developmental models have played a substantial role in scholarly explanations of the world and in the guidance of public policy. Our research was motivated by the idea that these models have been disseminated around the world to ordinary people where they have dramatic effects on individual lives and social change. Testing the extent to which such models have penetrated to the grassroots internationally and influence behavior is an important agenda for sociology.

Our research was motivated by fifteen interrelated conceptual, substantive, and methodological questions concerning the ways in which individuals view international developmental hierarchies. We investigated these fifteen issues using panel data from a sample of Taiwanese college students.

Our data overwhelmingly suggest that Taiwanese college students are able to rate countries on their development levels (Question 1). Also, various countries are rated differently and arrayed hierarchically according to their perceived development (Question 2). Our research also provides exceptionally strong evidence about the high stability of the aggregate level perceptions of the developmental hierarchy (Question 3). The average ratings of country development across three waves of observation reveal the existence of an exceptionally stable

aggregated perception of developmental hierarchy among this group of college students. In addition, even the detailed distributions of aggregated ratings for country development are distributed exceptionally consistently across different waves (Question 4). Also, our data indicate that, on average, Taiwanese students rate countries on development very similarly to the UN Human Development Index (Questions 12 and 13). Furthermore, large fractions of individual students have developmental hierarchies that closely match the UN HDI (Question 14).   These results provide strong support that the ideas of development and development hierarchies are widespread in Taiwan, at least among college students, and that these Taiwanese perceptions closely match the development ideas existing in world culture.

Our data also suggest that there are individual differences in the ways Taiwanese college students use our eleven-point scale in rating countries on development (Question 9). Some students tended to rate all countries high while other students tended to rate all countries low. Also, some students tended to use a relatively small part of our eleven-point rating scale, while others spread their country ratings over a larger part of the scale. These differences in response styles are also measured quite reliably and are very stable across time (Question 10).

It is important to note that this interpretation of individual differences in means and standard deviations of development rating scores assumes that they reflect personal rating styles. Another possible interpretation is that these individual differences reflect real substantive differences between respondents and not just response styles. That is, some respondents may have more positive (or more optimistic) views than others of development levels in the world. Also, some respondents may have more varied ratings of development than others. It will require additional research to decide between these interpretations.

We used the technique of structural equation modeling to simultaneously estimate the reliability and over-time stability of individual perceptions of country development. This study provides considerable evidence that the individual perceptions of the developmental levels of countries can be measured reliably (Question 5). Our data show that substantial fractions of the variance in individual perceptions of country development are commonly shared with the underlying concept. This holds for both the raw country rating scores and the scores adjusted for response styles (Questions 5 and 11). It also holds for the individual correlations between UN HDI scores and individual country ratings (Question 15).

Furthermore, our results indicate that the reliabilities of the individual perceptions of country development compare favorably to the reliabilities of attitudinal variables that are frequently used in social research (Question 7). This holds true for measures included in the Taiwan student panel survey and for measures used in many other social science surveys. Although there is room for improvement in eliminating measurement error, the measures for perceptions of country development we employed have similar capability to support high quality empirical research as standard attitudinal variables used in many social science studies.

We noted earlier that the aggregate stability of country ratings across time is exceptionally high. This is, of course, partially due to compensating differences from year to year, but our structural equation analysis suggests that there is also considerable stability at the individual level. This is true for both the raw and adjusted scores (Questions 6 and 11). It is also true for the individual correlations between respondent country ratings and UN HDI scores for the same countries (Question 15). The cross-wave stabilities of developmental ratings are also very comparable to the cross-wave stabilities of other attitudinal variables (Question 8). These results indicate that individual respondents' ratings of country development are very stable and are not simply the products of individual transient thoughts.

Thus, our research provides considerable support for the expectation that the ideas of development and modernization have permeated into the grassroots. Furthermore, this research indicates that the ideas that people have about developmental hierarchies can be measured reliably at both the aggregate and individual levels. Such ideas also have considerable stability at both the aggregate and individual levels. These findings provide a foundation for future research concerning the perceptions of ordinary citizens concerning development and development hierarchies.

These results, of course, cannot be generalized beyond the population of students attending one university in Taiwan. We expect that our results are not limited to this population, but further research will be needed to see if our results hold for students in other Taiwanese colleges, for the general population of Taiwan, or for other places around the world. We close this paper with a call for research to see how our results are similar or different in other places.

References

Abu-Lughod L. Remaking Women: Feminism and Modernity in the Middle East. Princeton, NJ: Princeton University Press; 1998.

Ahearn LM. Invitations to love: Literacy, love letters, and social change in Nepal. Ann Arbor, MI: University of Michigan Press; 2001.

Alwin DF. Aging, personality, and social change: The stability of individual differences over the adult life span. Featherman DL, Lerner RM, Perlmutter M, editors. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Publishers; 1994.

Alwin DF. The margins of error: A study of reliability in survey measurement. Hoboken, NJ: Wiley-Blackwell; 2007.

Alwin DF. Problems in the estimation and interpretation of the reliability of survey data. Quality and Quantity. 1989;23:277-331.

Amin S. Eurocentrism. New York: Monthly Review Press; 1989.

Baker DP, LeTendre GK. National Differences, Global Similaries: World Culture and the Future of Schooling. Stanford: Stanford University Press; 2005.

Baker LD. From Savage to Negro: Anthropology and the Construction of Race, 1896-1954. Berkeley and Los Angeles: University of California Press.; 1998.

Barrett D, Frank DJ. Population control for national development: From world discourse to national policies. In: Boli J, Thomas GM, editors. Constructing world culture: International nongovernmental organizations since 1875. Stanford, CA: Stanford University Press; 1999. p. 198-221.

Benavot A, et al. Knowledge for the Masses: World Models and National Curricula, 1920-1986. American Sociological Association. 1991;56:85-100.

Binstock G, Thornton A. Knowledge and use of developmental thinking about societies and families among teenagers in Argentina. Demografia. 2007;50:75-104.

Boas F. Race, language and culture. New York,: The Macmillan company; 1940.

Bock KE. The Acceptance of Histories: Toward a Perspective for Social Science. Berkeley, CA: University of California Press; 1956.

Böröcz J. The Fox and the Raven: The European Union and Hungary renegotiate the margins of 'Europe'. Comp Stud Soc Hist. 2000;42:847-75.

Böröcz J. Goodness is Elsewhere: The Rule of European Difference. Comp Stud Soc Hist. 2006;48:110-38.

Boyle EH, editor. Female Genital Cutting: Cultural Conflict in the Global Community. Baltimore: The Johns Hopkins University Press; 2002.

Caldwell JC, et al. The causes of demographic change: Experimental research in South India. Madison, WI: The University of Wisconsin Press; 1988.

Chabbott C, editor. Constructing Education for Development: International Organizations and Education for All. New York: Routledge Falmer; 2003.

Chakrabarty D, editor. Provincializing Europe, Postcolonial Thought and Historical Difference. Princeton: Princeton University Press; 2000.

Coenders G, et al. Stability of Three-Wave Simplex Estimates of Reliability. Structural Equation Modeling-a Multidisciplinary Journal. 1999;6:135-57. doi: Doi 10.1080/10705519909540125.

Comaroff J, Comaroff J, editors. Ethnography and the Historical Imagination. Boulder: Westview Press; 1992.

Cronbach LJ. Further Evidence on Response Sets and Test Design. Educ Psychol Meas. 1950;10:3-31.

Cunningham WJ, et al. The Ipsative Process to Reduce Response Set Bias. The Public Opinion Quarterly. 1977;41:379-84.

Dahl G, Rabo A. Kam-ap or take-off: Local notions of development. Stockholm: Stockholm Studies in Social Anthropology; 1992.

Drori GS, Krücken G. World society: A theory and a research program in context. In: Krücken G, Drori GS, editors. World society: The writings of John W. Meyer. Oxford: Oxford University Press, USA; 2009.

Ferguson J, editor. Expectations of Modernity: Myths and Meanings of Urban Life on the Zambian Copper Belt. Berkeley, CA: University of California Press; 1999.

Frank DJ, et al. Worldwide Trends in the Criminal Regulation of Sex, 1945 to 2005. Am Sociol Rev. 2010;75:867-93.

Guneratne A. Shaping the Tourist's Gaze: Representing Ethnic Differences in a Nepali Village. The Journal of the Royal Anthropological Institute. 2001;7:527-43.

Haq Mu. Reflections on Human Development. New York: Oxford University Press; 1995.

Harris M. The rise of anthropological theory. New York, NY: Crowell Company; 1968.

Heise DR. Separating reliability and stability in test-retest correlations. Am Sociol Rev. 1969;34:93-101.

Hodgen MT. Early anthropology in the sixteenth and seventeenth centuries. Philadelphia, PA: University of Pennsylvania Press; 1964.

International Monetary Fund. World Economic Outlook: Sustaining the Recovery. Washington, D.C.2009.

Jennings F, editor. The invasion of America : Indians, colonialism, and the cant of conquest. Chapel Hill, NC: University of North Carolina Press; 1975.

Jöreskog KG. Analyzing psychological data by structural analysis of covariance matrices. In: Kranz DH, Atkinson, R. C., Luce, R. D. and Suppes, P., editor. Measurement, psychophysics, and neural information processing. San Francisco: Freeman; 1974. p. 1-56.

Jöreskog KG. Estimation and testing of simplex models. British Journal of Mathametical and Statistical Psychology. 1970;23:121-45.

Justice J, editor. Politics, Plans, and People: Culture and Health Development in Nepal: University of California Press; 1986.

Liechty M. Suitably Modern: Making Middle-Class Culture in a New Consumer Society. Princeton, New Jersey and Woodstock, Oxfordshire United Kingdom: Princeton University Press; 2003.

Mandelbaum M. History, man, and reason: A study in nineteenth-century thought. Baltimore/London: The Johns Hopkins Press; 1971.

Melegh A. On the East-West Slope Globalization, nationalism, racism and discourses on Central and Eastern Europe Budagest: Central European University Press; 2006.

Melegh A, et al. Perceptions of Societal Developmental Hierarchies in Europe and Beyond: A Bulgarian Perspective. European Sociological Review. 2012;0:1-13. doi: 10.1093/esr/jcs010.

Meyer JW, et al. World society and the nation-state. Amer J Sociol. 1997;103:144-81.

Miles I. The New Post-Industrial State. Futures. 1985;17:588-617.

Nisbet RA. Social Change and History. New York: Oxford University Press; 1975/1969.

Osella F, Osella C. Once upon a time in the West: Stories of migration and modernity from Kerala, South India. Journal of the Royal Anthropological Institute. 2006;12:569-88.

Pigg SL. The Credible and the Credulous: The Question of 'Villager's Beliefs' in Nepal. Cult Anthrop. 1996;11:160-201.

Sanderson SK. Social evolutionism: A critical history. Oxford: Basil Blackwell; 1990.

Sumner A, Tribe MA. International development studies : theories and methods in research and practice. Los Angeles ; London: SAGE; 2008.

Szreter S. The idea of demographic transition and the study of fertility change. Population Devel Rev. 1993;19:659-702.

Thornton A. The developmental paradigm, reading history sideways, and family change. Demography. 2001;38:449-65.

Thornton A. Reading history sideways: The fallacy and enduring impact of the developmental paradigm on family life. Chicago, IL: University of Chicago Press; 2005.

Thornton A, et al. Knowledge and Beliefs about National Development and Developmental Hierarchies: the Viewpoints of Ordinary People in Thirteen Countries. Soc Sci Res. 2012;41:1053-68.

Tilly C. Big structuring, large processes, huge comparisons. New York: Russell Sage Foundation; 1984.

Todorova M, editor. Imagining the Balkans. New York: Oxford University Press; 1997.

Tsutsui K, Wotipka CM. Global civil society and the international human rights movement: Citizen participation in human rights international nongovernmental organizations. Soc Forces. 2004;83:587-620.

United Nations Development Programme. Human Development Report 2009: Overcoming Barriers, Human Mobility and Development. New York: United Nations; 2009.

United Nations Development Programme. Human Development Report 2010: 20th Anniversary Edition: The Real Wealth of Nations: Pathways to Human Development. New York: United Nations, 2010.

Wallerstein I, editor. Unthinking Social Science: The Limits of Nineteenth Century Paradigms. Cambridge: Polity; 1991.

Wang Z. Women in the Chinese Enlightenment: Oral and Textual Histories. Berkeley: University of California Press; 1999.

Wolff L, editor. Inventing Eastern Europe: The Map of Civilization on the Mind of Enlightenment. Stanford, California: Stanford University Press; 1994.

World Bank. Data and Statistics: A Short History. Cited from the following Website on January 27, 2010: web.worldbank.org.2010.

Xie Y, et al. Societal Projection: Beliefs Concerning the Relationship between Development and Inequality and Inequality in China. Soc Sci Res. 2012;41:1069-84.