Creating Questions and Protocols for an International Study of Ideas about Development, Demographic Behavior, and Family Life

Arland Thornton – University of Michigan
Alexandra Achen – University of Michigan
Jennifer Barber – University of Michigan
Georgina Binstock – Centro de Estudios de Población, Buenos Aires
Wade Garrison – University of Michigan
Dirgha Ghimire – University of Michigan
Wang Guangzhou – Chinese Academy of Social Sciences
Ronald Inglehart -- University of Michigan
Rukmalie Jayakody – Pennsylvania State University
Yang Jiang – University of Michigan
Julie de Jong – University of Michigan
Katherine King – University of Michigan
Ron Lesthaeghe – University of Michigan
Sohair Mehanna -- American University in Cairo
Colter Mitchell – University of Michigan
Mansoor Moaddel – Eastern Michigan University
Mary Beth Ofstedal – University of Michigan
Norbert Schwarz -- University of Michigan
Yu Xie -- University of Michigan
Li-shou Yang -- University of Michigan
Linda Young-DeMarco – University of Michigan
Kathryn Yount – Emory University

September 2008

Abstract

This paper reports work to understand people's knowledge and perceptions of social and economic development and the ways in which they relate to family and demographic behavior. This project was motivated by recent work suggesting that beliefs and values concerning development and related matters are important influences on demographic and family behavior. Our ultimate substantive goal was to study the consequences of holding specific viewpoints about development for several family and demographic behaviors, including marriage, childbearing, migration, and maternal and child health. When we began our work, we knew of no existing tools to measure how people understand and evaluate the ideas of development. Thus, our challenge was to create and evaluate measures of people's knowledge of and views about development that could be used in surveys in diverse settings around the world. Our research efforts to date have included data collection using in-depth interviews, focus groups, cognitive interviews, and surveys in eleven places around the world: Albania, Argentina, China, Egypt, Iran, Iraq, Nepal, Saudi Arabia, Taiwan, the United States, and Vietnam. In this paper, we describe our approach to create and test questionnaires and protocols for measuring ideas and beliefs about development in a variety of countries. We discuss specific problems that we encountered, along with lessons we have learned. We also provide preliminary evidence of our success in measuring developmental concepts and discuss the implications of our experience for other international data collections.

Acknowledgement

Creating Questions and Protocols for an International Study of Ideas about Development, Demographic Behavior, and Family Life

1. Introduction

This paper reports work undertaken to understand people's knowledge and perceptions of social and economic development and the ways in which they relate to family and demographic behavior. This project was motivated by recent work suggesting that beliefs and values concerning development and related matters are important influences on demographic and family behavior (Thornton 2005). Our substantive goals in this project are threefold: 1) to estimate the extent to which ideas of social and economic development are known and accepted or rejected; 2) to evaluate the factors influencing ordinary people's knowledge and acceptance or rejection of developmental ideas; and 3) to study the consequences of holding specific viewpoints about development for a wide range of behaviors, including marriage, childbearing, migration, and maternal and child health. We want to evaluate these substantive questions in a wide array of different countries.

When we began our work, we knew of no existing tools to measure how people in everyday life in even one country understand and evaluate the ideas of development. Thus, our challenge was to create and evaluate measures of people's knowledge of and views about development that could be used in surveys in many diverse settings around the world. Our research efforts to date have included data collection using in-depth interviews, focus groups, cognitive interviewing, and surveys in eleven places around the world: Albania, Argentina, China, Egypt, Iran, Iraq, Nepal, Saudi Arabia, Taiwan, the United States, and Vietnam.

In this paper, we describe our approach to create and test questionnaires and protocols for measuring ideas and beliefs about development in a variety of countries. In Section 2, we specify the developmental model and its basic propositions about social change, family, and demography. In Section 3 we describe our organizational approach and initial steps in designing sequentially projects in several countries without concern for cross-cultural comparability. Section 4 explains how we used the experience and knowledge accumulated from our work in individual countries to prepare questionnaires and protocols for use in international comparative projects. In Section 5 we discuss specific problems that we encountered in our project, along with lessons we have learned. Section 6 provides preliminary evidence of our success in measuring several aspects of developmental thinking. Finally, Section 7 discusses the implications of our experience for other researchers who are designing international data collections.

2. Developmental Concepts and Theories

An essential early step in the design of questions was to break down the complex and abstract ideas of development into concepts that are understood and used in everyday life. This step was necessary in order to create relatively simple questions that addressed meaningful concepts to people from many cultural and socioeconomic contexts.

We began with the developmental paradigm, a model of social change that posits that all societies progress through the same universal stages of development[1] (Burrow 1981; Harris 1968; Stocking 1968, 1987; Nisbet 1969; Smith 1973; Sanderson 1990; Mandelbaum 1971; Thornton 2001, 2005). People using this paradigm believe that, at any one point in time, there is a hierarchy of countries on a developmental ladder, that is,

---

[1] We do not present these developmental concepts as true and good but as ideas that can be important for people in everyday life whether or not they are true or good.

a scale of nations. Advocates of this paradigm suggest that the most developed or modern societies are in Northwest Europe and the Northwest European diaspora, while other societies are less developed, developing, or traditional. The United Nations today ranks countries on a numerical index of Human Development (HDI). One of our research goals was, therefore, to learn whether ordinary people know about and endorse this model of a developmental ladder on which societies are ranked.

Because many western scholars commonly placed Northwest Europe and the Northwest European diaspora at the pinnacle of civilization (Millar 1979/1779, Mill 1848, Guizot, 1890; Tylor 1871), certain attributes that existed in the West such as industrial economic organization, urban living, high levels of education, technological skill, high levels of consumption, and low levels of mortality came to define in the West what it meant to be modern or developed. A second research goal, therefore, was to ascertain whether and to what extent ordinary people around the world associate the same societal attributes with development.

For centuries it has been understood that while there was considerable heterogeneity in family patterns both inside and outside Northwest Europe, certain family patterns in Northwest Europe differed greatly from those in many other parts of the world. These distinguishing family attributes of Northwest Europe became associated with modernity or development whereas the family attributes of other places were characterized as traditional (Thornton 2001, 2005). Among the family attributes that Western observers said characterized places outside of Northwest Europe and defined traditional families more generally were the organization of activities around the family, extensive family solidarity, great respect for the elderly, large and complex households,

young age at marriage, arranged marriage, a low status of women, and high and uncontrolled fertility. In contrast, the family attributes that for some Western scholars came to characterize Northwest Europe also came to be associated with development and modernity. These attributes included the organization of many activities outside the family, individualism, limited respect for the elderly, nuclear households, an older age at marriage, affection in the mate selection process, a higher status of women, and low and controlled fertility. The frequent use of these definitions of traditional and modern families led our team to want to investigate whether ordinary people around the world use similar definitions.

At the same time, we recognized that in some cases models for family and social life long held outside the West may overlap with the Western models described above. Indeed, intellectual and religious elites in some settings outside the West have viewed Western models of development as culturally imperialistic and have sought to claim as indigenous certain desirable attributes of modern family by locating their guiding principles in historical religious texts (e.g., Hoodfar, 2008, forthcoming; Yount and Rashad, 2008, forthcoming).

Western scholars not only have observed the correlations between the societal and the familial and demographic attributes of a society, but also have created theories about causal connections between the Northwest European family system and the Northwest European social and economic system (Thornton 2001, 2005). Most of these scholars saw this causation as being the effect of socioeconomic development on family and demographic change, but others hypothesized an effect of family and demographic change on development. We therefore aimed to evaluate the extent to which individuals

in everyday life believe that modern societal attributes causally affect modern familial and demographic attributes and that modern family and demographic attributes causally influence modern societal characteristics.

Developmental ideas provide a framework not only for how development works but also for evaluating the value of certain societal and familial traits. In this framework, the attributes of family and society that are specified as modern also are perceived to be desirable. The framework also suggests that the good things of development can be attained through discipline and hard work. Consequently, we wished to ascertain the extent to which people evaluate certain aspects of modern family and societal life as being better or worse than traditional dimensions of family and society and the extent to which people view the good things as attainable.

3. Project Organization and Country Specific Work

As discussed by Pennell and colleagues (forthcoming), the organization and strategy of international research projects depends upon many factors, including the source and flow of funds, the experience and knowledge of the research team, and the availability of collaborators with the necessary interests and research infrastructure. The organization of our project was also influenced by the fact that our aim was to measure people's knowledge of and adherence to a complex and sophisticated set of beliefs and worldviews while most surveys focus on methods appropriate for measuring experience, behavior, and attitudes. In the beginning we also had only a very small team of researchers and very limited resources.

Several decisions grew out of these considerations. We decided that, to the extent possible, we would utilize multiple methodological approaches in order to obtain wide

and diverse perspectives on the use of developmental thinking. We also decided to start the project by utilizing the methodological expertise that was available to us in our research team. We also knew that we would need to start small in order to build up the necessary expertise and knowledge to serve as the foundation for a more ambitious international comparative project. This led us to decide to create measurement procedures, protocols, and questions that could be used in one country without any explicit concern for international comparability. Once we had, at least partially, accomplished our goals in one country, we could apply what we had learned in that country to other countries and repeat the process of creating and testing measures, with modifications as necessary. Given our starting point, this sequential approach seemed more desirable in the beginning than a parallel or simultaneous approach in several countries, although, as described below, we later used a simultaneous approach[2].

The work on the project began when a small team of sociologists and demographers with experience and infrastructure in Nepal initiated a small mixed-methods study in that country. This team had expertise and experience in exploratory interviewing, focus groups, and survey research and decided to launch fieldwork in 2003 with informal exploratory discussions with residents in Chitwan Valley. These interviews provided insights into how ordinary Nepalis think about development and the factors associated with it. Guided by the insights from these informal discussions, we next conducted semi-structured interviews and focus group discussions about

---

[2] For a discussion of sequential, parallel, and simultaneous approaches to the design of international comparative questions, see Harkness et al 2003.

developmental thinking and family life. Following that, we administered in Nepal a questionnaire with a sample of adults using face-to-face interviews[3].

With the Nepal experience, we were ready in 2003-2004 to launch a mixed-methods project in Argentina. Given the expertise of our Argentina team in focus groups, we designed a project to conduct focus groups among high school students in the city of Buenos Aires and in rural schools in northern Santa Fe Province. Immediately prior to the focus groups we asked each of the focus group participants to complete a self-administered survey questionnaire. This questionnaire also was administered to additional students in the same Buenos Aires high schools and in additional rural high schools in northern Santa Fe[4].

Following the mixed method project in Argentina, we designed a similar project in 2006 with adults in Cairo, Egypt. As in Argentina, this project conducted interviews with the focus group participants immediately before their participation in the focus groups. Unlike in Argentina where the questionnaires were filled out by respondents, in Egypt the interview data were collected in semi-structured interviews using an explicit sequence of open- and closed-ended questions[5].

Another exploratory project was designed for the United States where, in 2006, we conducted a set of cognitive interviews using a questionnaire that included individual survey questions and probes asking respondents to explain their answers. In addition, we designed a set of questions that were fielded as a supplement to the University of

---

[3] This fieldwork in Nepal was directed by Dirgha Ghimire.
[4] This fieldwork in Argentina was directed by Georgina Binstock.
[5] This fieldwork in Egypt was directed by Kathryn Yount and Sohair Mehanna.

Michigan Survey of Consumers—a national telephone survey of adults. This study also included experiments to evaluate the effects of question wording and question ordering[6].

As we gained more experience and as more researchers became interested in our substantive research goals, additional research opportunities presented themselves. These consisted primarily of adding small sets of our developmental questions into studies conducted for other purposes. In each case our goal was to collect country-specific data without trying to make the data comparable across countries.

Our first such experience was in Vietnam where we added a small battery of questions about development and family life to the 2003 Red River Delta Family Survey. An innovation in Vietnam was that we randomly probed a sub-sample of respondents, asking open-ended questions about what they thought modernization was, and what the positive and negative aspects of modernization were[7]. In the fall of 2004, we added a small module of questions to a study about democratic ideas and values among college students in Taiwan using self-administered questionnaires. Inasmuch as this project in Taiwan annually collected data from a panel, it provided the opportunity to compare the answers of people over time[8]. In 2004 and 2006 in Iraq and in 2005 in Iran we added a small battery of our development and family questions to World Values Surveys being conducted among national samples in these countries[9]. In 2005, questions concerning developmental thinking were incorporated into a national UNICEF Survey on maternal

---

[6] This fieldwork was directed by Arland Thornton, Linda Young-DeMarco, and Jim Lepkowski.
[7] This data collection was directed by Rukmalie Jayakody.

[8] This data collection was directed by Li-Shou Yang.
[9] This data collection in Iran and Iraq was directed by Mansoor Moaddel.

and child health in Albania[10]. In 2005, surveys of young adults were conducted in three large cities each in Egypt and Saudi Arabia. In both the Egyptian and Saudi youth surveys our questions about development and family were added to questionnaires focused on politics, religion, and gender[11]. Finally, in 2006 we added a modest number of items about development, family, and inequality to a Family Policy Survey conducted in several Chinese provinces[12].

4. Designing Comparative International Protocols and Questions

As we accumulated information and insights from our country-specific projects, we turned to the creation of questions that could be asked any place in the world. These questions would be useful for researchers working in an individual country as well as for people interested in international comparative work. In this comparative part of the project we had two sub-goals: first, creating a large battery of questions appropriate for measuring a wide range of dimensions of developmental thinking and its relationship to various aspects of family, demography, and social life; and, second, selecting from these questions a subset that could be asked in actual surveys in five countries.

An essential element for accomplishment of this task was a team of experts in relevant disciplines who brought wide cultural diversity and expertise. Our project drew together scholars in the fields of anthropology, demography, political science, psychology, and sociology. The team members also brought expertise from a range of epistemological approaches, including ethnography and survey research. The team members had knowledge and experience in countries as diverse as Albania, Argentina, Belgium, China, Egypt, Iran, Nepal, the United States, and Vietnam.

---

[10] The developmental idealism data were collected under the direction of Arjan Gjonca.
[11] This fieldwork in Egypt and Saudi Arabia was directed by Mansoor Moaddel.
[12] This data collection was directed by Wang Guangzhou

When we designed questions for use in multiple countries, we followed a simultaneous or de-centered approach to the extent possible by focusing on many countries simultaneously in designing our questions. That is, instead of designing our questions for one specific country and then generalizing those questions to other countries, we simultaneously designed our questions for use in many countries. In this data design work we included specific focus on six different countries—Argentina, China, Egypt, Iran, Nepal, and the United States—but also drew from the knowledge of our team about other places. Our de-centered international approach required that we work inductively and pragmatically in order to uncover the specific ways in which various ideas or concepts were understood in particular cultures.

We have tried to make translation a team project and an integral part of the questionnaire design process. The questions for our comparative project were written with team members including multilingual speakers who were both knowledgeable of our research concepts and goals and who were expert in at least two languages and cultures. Our questions were initially formulated in English, our lingua franca, but were designed with many countries in mind—a task that was facilitated both by having team members from many societies and by the fact that we had previous experience with similar questions in many different settings. In addition, the translation of questionnaires into other languages has been done iteratively, with the translations revealing changes required in the English version to facilitate comparability across languages. Multiple and coordinated iterations are required when the number of languages increases beyond two. It was sometimes necessary to make sequential adaptations in the final stages of implementation.

We created a large battery of questions because we wanted a compendium of questions that researchers could choose from for their various purposes. It turned out that the number of questions we included in the compendium is large enough to fill several reasonably-sized questionnaires. Our plan is to make this compendium of developmental questions available to the general research community.

We then selected a subset of questions from the compendium to include in a questionnaire appropriate for administering in a wide array of countries. This questionnaire has, to date, been fielded in surveys in representative samples of respondents in Argentina, China, Egypt, Iran, and the United States, with the fieldwork conducted in 2007 and 2008. Because of severe budget constraints and different methodological constraints in the various settings, we have followed different sampling and interviewing strategies in the five countries.

The study in Argentina, was conducted using face-to-face interviews with a national sample of adults living in urban settings, and the Chinese data collection was conducted using face-to-face interviews with adults living in Gansu Province. The Egyptian data collection was conducted in one district each in Qaliubia Governorate and in Fayoum Governorate using face-to-face interviews with a sample of adult women and their husbands, while the survey in Iran was conducted with face-to-face interviews with adult women in the city of Yazd. The U. S. data collection consists of two separate 15 minute supplements (in May and November 2007) appended to the Survey of Consumers, a nationally representative monthly telephone survey of American adults[13].These

---

[13] The Argentina, China, Egypt, Iran, and U. S. surveys were conducted respectively under the direction of Georgina Binstock, Wang Guangzhou, Kathryn Yount and Sohair Mehanna, Jalal Abbasi-Shavazi and Abbas Askari-Nodoushan, and Arland Thornton and Linda Young-DeMarco.

differences in samples and interviewing modes mean that strict comparability across

settings is not possible with these data[14]. Even with these sample restrictions, however,

we still can obtain crude comparisons of distributions across settings. Although each of

these five data collections can be used to infer to their respective sampling universes, we

refer to them as pilot studies in the context of our comparative international focus.

Because the final questionnaire was intended to be universally applicable, deciding upon

the set of questions to include was a challenging task. One of the things that became

apparent early in this process was that some of our theoretical concepts were not equally

applicable in all places, could not be asked in the same way in all countries, or were

considered too sensitive in some settings. In addition, some of our investigators had

more interest in some theoretical concepts than in others. These considerations led us to

divide the questionnaire into a core part that would be included in all country surveys and

an optional part that could vary across countries (see Harkness et al. 2003 and Smith

2003 for general discussions of this issue). Our goal has been for the core section to be

about 30-40 minutes in length. We have delegated the construction of the optional part of

the questionnaire to the people in charge of the data collection in a specific country.

5. Problems Encountered and Lessons Learned.

We now turn to a discussion of some of the problems encountered and lessons

learned in our country-specific and comparative international work.

5.1 Conceptual Coverage of the Development Concept

Our studies have shown that the concept of development is widely understood

around the world. In all of our study countries we have found one or more phrases that

---

[14] Although each of the five data collections interviewed adults, they used different age cut-offs for the adult population.

are very similar to the concept of developed or modern in English[15]. In many places the concept of development is frequently used in everyday discussions. The substantive implication is that the idea of development has been disseminated widely.

In addition, our work indicates that there is considerable overlap in the meaning of development in many different countries. It is widely understood that development is strongly related to socioeconomic factors. For example, in open-ended questions about the meaning of modernization, most respondents in Vietnam defined modernization in economic terms, citing advantages such as having enough to eat and having a good standard of living. Similarly, structured interviews and focus groups in Egypt revealed that development there meant things such as education, science and technology, a sound economy, job opportunities, and high quality and accessible medical services.

Our research also suggests that the definition of development used in many places is very similar to the one used by the United Nations in its Human Development Index (HDI), which is a composite of income, education, literacy, and health. We ascertained this by asking survey respondents to rate several countries on their level of development using an eleven-point scale from zero to ten. Inasmuch as the UN HDI also rates countries on development, it is possible to compare respondent ratings with the external HDI. Our preliminary analyses indicate that, on average, respondents in our surveys rate countries very similarly to the UN. Methodologically, this indicates that people in many places around the world have a common concept of development which can be investigated in surveys.

5.2 Variability of Geographical Location and Context

---

[15] We list here the phrases that we have used in some of the languages to denote the English concept of developed. In Nepali it is "bikas"; in Spanish it is "desarrollado"; in Arabic it is "tanmiya" and "takadum" (both words); and in Vietnamese it is "hiện đại hóa" (referring to both modernization and development).

When we asked respondents in several surveys to rate countries on their level of development, we learned quickly and not surprisingly that the countries known by respondents varies tremendously across the world. We handled this in our country-specific work by tailoring the countries rated to the location of the respondents.

However, when we extended this methodology to our comparative international work, we believed that it was useful to standardize the countries rated. This required us to choose a set of countries to be rated that were both relatively well known around the world and that represented a range of scores on the U.N. HDI. Among the countries that were pre-tested and later dropped were Sweden, Somalia, and Zimbabwe because many respondents in some countries were not familiar with them. Our list ultimately included Japan, Nigeria, India, the United States, China, Central African Republic, France, Brazil, and Pakistan. For people saying they did not know the Central African Republic, interviewers were instructed to tell them that it was a country in the middle of Africa. For the surveys that were conducted in countries that were not included on the standard list, we added that country as a tenth country to be rated.

5.3 Evaluation of Development

A consistent theme of the results of our country-specific work was the evaluative element of ideas about development. Our research suggests that socioeconomic development often is evaluated as a positive thing—as in the proposition that a modern society and economy are good and should be striven for. Socioeconomic aspects of development that were consistently seen as good included education, a high standard of living, and excellent health.

At the same time, the application of the terms modernity and development to family matters brings mixed associations and evaluations among respondents in some countries. Some Egyptians, for example, equated certain attributes of modern family with Western cultures, and therefore perceived these attributes as negative and flawed. Such attributes included excessive individualism, materialism, secularism, and a morality based on self- interest and questionable behavior, especially with regard to sexual relations. In Vietnam, negative aspects of modernization also centered on family behavior, especially increases in divorce and declines in obligation towards elderly parents. Both religion and political ideology can be sources of acceptance or rejection of various attributes of family life.

We expect that the evaluation of family and demographic attributes defined as developed or modern varies substantially across countries and groups within countries. In order to measure this variation, we have included in our comparative international surveys a set of questions designed to ask respondents to compare the desirability of certain family attributes. For example, we ask respondents to choose between two different ages at marriage and between two different kinds of living arrangements as better for most people. A related section of our comparative questionnaire asks respondents to evaluate whether certain attributes—such as age at marriage or divorce—will increase or decrease in the future. We also ask whether the respondent believes that the expected trend is a good thing or bad thing.

5.4 Competing Frameworks

Just as competing evaluations of the merits of family and demographic attributes emerged from our country-specific work, so did competing ideas about the causal

mechanisms underlying the evolution of traditional families into modern ones. Although many respondents saw structural changes such as increasing education and wage employment as drivers of family change, others emphasized ideational forces. Noteworthy here was the extent that teenagers in Argentina adhered to ideational theories of family change, with a significant number of participants indicating that changes in values and attitudes were the most important factors influencing family change (Binstock and Thornton, 2008). This insight led us to design questions asking people to tell us their views about the influence of the media, especially television, on family change.

5.5 Cultural context and perceived question meaning

One important lesson that emerged from our work in Argentina concerned the importance of cultural context and perceived question meaning for questions asking about the causes and effects of development. A number of questions that had worked well in Nepal about how development might change family life in Nepal and how particular family changes might affect the Nepalese economy and standard of living were initially adapted for Argentina. We did this by asking respondents in Argentina how development might change family life in Argentina and how family changes might affect the Argentinian economy. However, the responses provided suggested that the questions were perceived differently in Argentina than in Nepal because respondents were being asked to respond about changes in two different countries—in one case, Argentina, and in the other case, Nepal. This made the questions incomparable across the two settings.

This prompted us to adopt a vignette approach in which we standardized the frame of reference for all respondents in all countries. We did so by asking respondents in all settings to consider a hypothetical country that has the attributes of being low-income,

agricultural, with limited education, and high mortality. We then ask respondents to give their views on how development would change family life in that hypothetical country and how certain family changes would influence development in that same hypothetical country. One disadvantage of this vignette approach--compared to asking about a specific country, for example the respondent's country—is that it makes the question hypothetical and, perhaps, more difficult to answer. This suggests that when international comparability is not a goal that it might be better to ask about the causes and consequences for a respondent's own country.

Vignettes have been widely and successfully used in the social and health sciences. For example, Sadana and colleagues (2001) have given respondents in multiple countries descriptions of people with identical health states and asked them to rate the level of health among the described people—thereby permitting the researchers to evaluate differences in rating modes across countries. In another use of vignettes, King and colleagues (2004, 2006) asked respondents to rate a series of fixed vignettes to learn about the ways in which different respondents used the response categories. Then, the researchers used this information about differential response to adjust the scores that respondents gave concerning their own situations and experience. In another use of vignettes, Correll et al (2007) evaluated the effects of race, gender, and parental status on people's evaluations of job applicants. They did so by giving respondents descriptions of job applicants who had identical attributes and experiences except for race, gender, and parental status which were experimentally manipulated to see their effects on respondent evaluations.

Another example of the difficulty of constructing universal questions for all countries concerned the concept of attainability. A related concept to attainability is fatalism which refers to whether or not people believe that things are outside of their control. We found the concept of attainability to be very difficult to implement cross-culturally and decided that for many countries, it would be best to use culturally-specific questions about fatalistic ideas using locale-specific language and idioms. For these reasons we have relegated the section on attainability versus fatalism in our comparative international surveys to the optional rather than the core category of questions.

Another set of questions that we believed was important to include in our surveys that turned out to be difficult to implement cross-culturally concerned religion and religiosity—institutions that have historically provided alternatives to developmental models. This difficulty became apparent as we dealt with the complexity of religion in China and Vietnam. Historically there was not one definite term to describe the phenomenon labeled in the West as "religion". This complexity increased even more with the introduction of socialism into China and religion was defined to be the opposite of science and contradictory to socialism (Linshu 1994). Religion in China is composed of a combination of three religious traditions: ancestor worship, Buddism, and Taosim (Teiser 1995). Each of these religious traditions has its own saints or,gods, namely, deceased family members, Buddha(s), and locally worshipped God(s) (Tan 1983). Chinese religion is also more of a "diffused" religion, where practices are mixed with non-religious elements (Yang 1962; Kazuo 1995; Tan 1983), Our resolution to this issue was to make religion a country-specific item rather than part of the core questionnaire across all countries.

5.6 Respondent Knowledge

One of the central issues that arose in our country-specific studies centered on the place of respondent's personal knowledge of specific phenomenon. Many of the questions we wanted to ask differed greatly from the standard questions often found in surveys in that they tend to focus on belief systems and understanding of the world rather than on facts, experiences, behaviors, and attitudes. Thus, many of our questions violated the principle of survey research that respondents should have the necessary personal knowledge to answer a question. That is, many of our questions asked about things most respondents had very little direct experience with and things that were outside their immediate realm of knowledge. For example, we asked questions concerning development levels in places respondents had never even visited. We also asked respondents to compare family attributes between developed and not developed countries when their own experiences were probably limited to a small range of countries?

In the design work for our 2006 data collection in the United States, we attempted to make our questions more concrete by supplying respondents descriptive information about a country or about the meaning of "development". For example, in Nepal we had asked respondents to compare Nepal and the United States on several dimensions and we wanted Americans to make the same comparisons. One experiment was to tell American respondents that Nepal was a mountainous country located in Asia between China and India. However, this strategy encouraged respondents to base their answers on the descriptions rather than on their overall impressions. This occurred as many respondents placed their focus on mountainous countries, in general, or to Asia, or to China or India specifically.

However, informal discussions, structured interviews, and focus groups in several countries indicated that people in everyday life can readily express opinions about development and other countries, even if they lacked first-hand knowledge about them. Indeed, most of the people we have interviewed have ideas about developmental topics and about other countries and can discuss these things in interview settings, even if their knowledge was gained through second-hand channels.

5.7 Response Category Design

Pretesting in Nepal revealed that forced choice questions with more than two response categories were problematic in that country because they required considerable probing in order to obtain the detailed information they were intended to obtain. This format required multiple probes, lengthened the interview, and increased respondent burden. Questions using dichotomous answer options proved easier to administer and reduced response time and effort. They also made it easier to create response categories that are consistent across languages (Smith 2003). Also, dichotomous response categories have been found to provide greater reliability than items with multiple categories (Alwin 2007). The pros and cons of multiple response categories in international surveys have been raised by others, including Smith (2003).

The number of response categories also emerged as an issue in our 2006 US survey when we asked respondents to choose whether certain family and societal attributes were more common in developed or less developed places. The Nepalese research that revealed problems with multiple-category answer scales suggested that a forced choice between these two categories was the preferred format, especially because our eventual goal was international comparability. However, in order to assess the effect

of using dichotomous response categories and trichotomous response categories, a split ballot experiment was included in the US survey. One half of the sample was offered a middle category "about the same" and the other half was presented with a dichotomous forced choice.

Analyses of these data indicate that more respondents chose the middle category "about the same" when it was explicitly offered than volunteered it as a preferred choice when it was not offered This result is expected as respondents in surveys tend to avail themselves of the options that are offered. However, most respondents in either split sample chose one of the two opposing answer options. The different formats also did not affect the ratio of the number of respondents choosing one of the polar categories to the number choosing the other polar category. In addition, both the dichotomous and trichotomous versions produced approximately the same amount of missing data. These results suggest that the dichotomous approach is acceptable in the United States. These results, together with those from Nepal support using dichotomous response categories in our comparative data collections.

5. 8 Dealing with Limited Variance

Our work revealed a dilemma for our design and analysis. Our research is motivated by the desire to understand the distribution of beliefs and values concerning development and to analyze their correlations with other dimensions of life. Accomplishing both goals has proven to be difficult whenever elements we wish to evaluate are endorsed either by an overwhelming percentage of respondents or by only a very small percentage. For example, 91 percent of Nepalis said that increased use of contraceptives would make Nepal richer. Only 10 percent of Chinese respondents said

that women's status will decline as countries undergo economic development. For our current work we have given priority to the descriptive goal of estimating levels of beliefs and values. As we move towards estimating how various factors influence developmental beliefs and values and how these in turn influence behavior, we will refine questions to ensure sufficient within-country variance.

6. Preliminary Evidence about Success of Country-Specific Studies

We now turn to an evaluation of our success in measuring developmental ideas. Because our five country pilot studies have only recently been completed, our evaluation has to rely on data from our earlier country-specific studies. We give special emphasis to data from the following countries: Argentina, Egypt, Iran, Iraq, Nepal, Saudi Arabia, and the United States. Because our analysis of the country-specific data is on-going and because of limitations of space, we can provide only basic evidence about the quality of the data. Also, in some cases our analyses are based on small and purposive samples and need to be replicated with larger representative samples.

We utilize several criteria for evaluating our success. One criteria focuses on people's ability to use and apply developmental concepts in their responses. A lack of understanding and knowledge of developmental thinking would be revealed in respondents becoming frustrated, terminating the survey early, refusing to answer questions, saying that they do not know how to answer the questions, and providing answers that do not appear to be related to the questions. We also consider the comments of the interviewers and respondents concerning the interviews and examine the amount and type of item non-response, including breaking-off the survey. We also check for response patterns that suggest faulty interviewing or acquiescence as indicated by

respondents giving the same answers to different questions, including answers to similar questions that were worded in different directions. We also consider the reports of focus group moderators and survey interviewers and respondents evaluating experience in the various data collections.

We found substantial evidence across all of the studies that the questions about developmental ideas were generally well understood. Despite vast cultural, economic, and geographic differences, most respondents in each of the countries reacted positively to the questions.

Due to the complexity of developmental thinking, the surveys often contained several questions about knowledge and beliefs, with many being about relatively complex and abstract ideas. Considering that the surveys ranged from around 17 minutes (United States) to 70 minutes (Nepal) in length, it is impressive that for the majority of the surveys, no respondents terminated the survey early. In fact the exceptions to this rule were terminations unrelated to the developmental questions. For example, the three break-offs in Nepal were caused by the respondents being unable to communicate in Nepali, the national language. Apparently, the questions on developmental ideas were not so difficult as to lead to respondent frustration and termination of the interview.

Further evidence of the ability of the respondents to answer our developmental questions is the relatively low level of non-response. By illustration, item non-response for the questions asking respondents to relate family and societal attributes in various places rarely exceeded 3-5 percent of the responses, and often remained at or below 1%. The data from the self-administered questionnaires with high school students in Argentina exemplifies the lowest levels of missing data, with most of the questions

having only .5-1.5% missing data. Iran and Iraq, however, had examples of higher levels of missing data for these questions with 8-10%. Also note that the vast majority of the missing data are the result of respondents reporting that they don't know the answer to a question. Another important consideration, addressed in greater detail later, is that we found that the surveys that probed for a response after a "don't know" answer had substantially lower levels of missing data than surveys that did not probe after a "don't know" response.

One set of sections that did tend to have slightly higher levels of missing data were the scales that asked respondents to rate countries from 0 to 10 (or 1 to 10 in a few settings) on a particular characteristic such as development, income, gender equality, education, freedom, and morality. The higher missing data for such questions was primarily due to the fact that we often asked respondents to rate 5-10 different countries, and some people said that they did not know a country well enough to rate it on a particular characteristic. However, we found that in countries where probes were not used, there were substantially higher levels of "don't know" responses. Not surprisingly, we also found that respondents tended to reply don't know substantially less often when rating large countries and countries well known in the region and the world.

The responses of the 84 respondents in the questionnaires preceding the focus groups in the Cairo area illustrate the difficulties with these types of questions. In that study we asked respondents to rate countries from 1 to 10 on "economic development," "education," "income/wealth," and "gender equity". The percentage of non-response ranged from 1 (Japan and the United States) to 29 (Zimbabwe). Relatively high percentages of "don't know" responses on these scales also occurred for Nigeria,

Sweden, and Brazil. Again, the higher percentages of item non-response probably are due to the Egypt survey not probing after "don't know" responses. Also, in other Middle Eastern country surveys (Iran, Iraq, Egypt. Saudi Arabia) there were typically higher levels of "don't know" responses than in other countries (US, Nepal, Argentina) even before probes. This result may indicate a greater reluctance in these countries to answer these questions if the respondents do not think that they have enough information.

Probing is a good way to reduce the amount of non-response for these country rating questions. The standard probe that we have created for this purpose is, "Even if you don't know exactly, about where would you put [country X]?" One concern with the probing procedure is that the quality of the probed data would be lower than the quality of the non-probed data. Nevertheless we found that the distribution of responses after the probe was relatively similar to the distribution of responses given without a probe. For example, Nepalis had particular difficulty rating (on a 0-10 scale) the education and development of Somalia. Ten percent failed to provide an initial answer for Somalia, but after the probe only 4% failed to provide a response. However, the respondents who provided the score only after the probe rated Somalia only slightly lower than those who provided a response without the probe. For the other 9 countries we asked about in the Nepal data collection, only 0.9-2.2 percent of respondents failed to provide an estimate of development level after the follow-up probe of a don't know response. We also found the distributions of the pre and post probe responses were very similar in the US survey.

It is notable that despite some respondents having difficulty rating some countries on development, they still, on average, rated the countries very similarly to the ratings provided by the United Nations Human Development Index (UNDP 2003). This appears

27

to be true in all of the countries where we have asked respondents to rate countries on development.

We examined the quality of the responses in some countries by setting up methods to examine response patterns that were likely to reflect methodological artifacts rather than substantive information. For example, in Nepal we were able to ask some questions in the opposite direction of the rest of the questions in a section to see if respondents were using the same responses to oppositely-worded questions. Because respondents tended to notice the reverse coded questions in the set of questions, there is evidence of response set bias for only a small number of the respondents. This result provides further evidence that the questions we asked are understandable and the answers meaningful to the vast majority of respondents. It also suggests that the low amount of non-response for the various questions is not due to respondents being agreeable to every question asked, without providing a thought-out response.

Although we have to this point primarily discussed data from surveys responses, some of our strongest evidence concerning data quality comes from the several focus groups and in-depth interviews. During some focus groups in Egypt, for example, the discussions of what is modern and how that relates to family life were especially vibrant. Similar discussions in the US, Argentina, Nepal and other places suggested that despite some respondents not necessarily having a formal name for developmental ideas, they were able to respond to questions on the topic and rarely had trouble conveying their beliefs concerning such ideas. Our focus group moderators and survey interviewers in these countries also reported that almost all study participants seemed interested in the topic and willing and able to discuss developmental thinking. The interviewers,

moderators, and respondents also indicated that the focus group discussions and surveys about developmental thinking were engaging and enjoyable.

Despite the overall success of our surveys, different questions proved to be challenging for respondents in each country. One example is particularly informative. In our data collections in the Middle East we discovered that respondents made important distinctions between traits that they considered Western and traits that they considered modern.   This distinction has proven to be important in Middle Eastern and some non-Middle Eastern settings because respondents' evaluations of specific traits as good or bad appears to be based partially on whether they consider the traits to be Western. It is also possible that their positive or negative evaluation of a trait could influence whether or not they consider that trait to be Western. This finding has led us to work more intensely on trying to understand the ways in which people from different cultures attribute various traits.

7. **Conclusions**

In this paper we have presented our efforts to measure the ideas and beliefs that ordinary people around the world have about social and economic development and the ways they relate to family and demographic behavior. As we described earlier, socioeconomic development is a sophisticated set of concepts, beliefs, and values that include theories about many aspects of family, demographic, and societal change, along with a model for evaluating the relative merits of various dimensions of social life. Our research is motivated by our desire to know how widespread these models and beliefs are, the factors influencing people's acceptance or rejection of them, and the consequences of these models and beliefs for changing lives and social institutions.

When we began this project, we were unaware of any existing measures of these models and concepts that we could use in our research. We have spent considerable time and energy in formulating and evaluating new questions for this research program—a task made especially difficult because of our goal of designing questions appropriate for the many and diverse societies of the world.

We began this very substantial task with a small research team and a very small budget. This lack of general expertise and resources prevented us from immediately constructing questions and protocols for a large international comparative study. Instead, we started by working in a single country–Nepal, where we had access to a research infrastructure with which to conduct preliminary work. In this country, we started with a multi-method approach including informal interviews, semi-structured interviews, focus groups, and finally a survey interview with a sample of respondents. With this experience and the interest shown by other researchers, we extended our work to other countries—doing surveys in several countries, and in some places cognitive interviewing, focus groups, and experiments. Our work in these countries was cumulative in that we used information gleaned in one setting to inform the work in other settings, but in each case our desire was to design questions and protocols that would work in the specific country involved rather than to have materials that could be used comparatively across countries.

With these country-specific projects completed, we had available to us the accumulated wisdom from our experience in a substantial number of diverse settings. We had also gathered a multinational, multilingual, and interdisciplinary team with great experience in multiple settings around the world. With this team in place we undertook a

process of parallel and simultaneous construction of comparable questions and interviewing protocols for use across a wide variety of settings—translating questions iteratively as we proceeded.

We believe that this multifaceted approach to questionnaire design has been necessary and largely successful. Although our evaluation of our results is only in the beginning phases, our analyses reported above give considerable promise for our efforts being at least partially successful. This level of success, however, would have been exceptionally difficult—perhaps even impossible—had we gone directly to the preparation of comparative questions for many cultures around the world. In the early phases of our work our lack of country-specific experience and our lack of colleagues from many countries would have likely led such an effort to frustration and dissatisfaction. We recommend the general approach we used to other researchers facing a similarly complex research agenda.

Our experience also confirms the experience of many other cross-cultural researchers that creating comparable measures across very different societies with divergent languages is an exceptionally challenging undertaking. Although we believe that it is possible to measure many things comparably, perfect exactness is a very difficult standard for many concepts. And, for some concepts it is simply impossible. In addition, it is very easy in a project iterating questions and translations across several languages and cultures to declare a product finished before all of the iterations are completed. In addition, because of the large number of cultures and languages in the world, it is probably ultimately impossible to accomplish this task simultaneously for all people and places. This means that while a process like ours can work at least partially for a

specified set of countries, the extension of results to other places will ultimately require a less satisfactory sequential design approach.

Our future plans include continued evaluation of the data that we have collected. We expect that this analysis will shed important light on the areas where we have succeeded and failed in producing cross-culturally comparative measures and data. This analysis also is likely to give us guidance for improvement for future data collections. Although considerable room for improvement exists, we believe that we are currently well-positioned to design larger studies to document the distribution of developmental ideas, beliefs, and values around the world, and to analyze their determinants as well as their consequences.

Reference List

    1. Alwin, D. F. (2007). <u>Margins of Error: A Study of Reliability in Survey Measurement</u>. John Wiley and Sons.

    2. Binstock, G., & Thornton, A. (2006). Knowledge and use of developmental thinking about societies and families among teenagers in Argentina. Published as a Working Paper of the Population Studies Center, University of Michigan. Also forthcoming in <u>Demografia.</u>

    3. Burrow, J. W. (1981). <u>Evolution and society</u>. Cambridge: Cambridge University Press.

    4. Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? <u>American Journal of Sociology, 112</u>(5), 1297-1338.

    5. Guizot, F. (1890). <u>The History of Civilization</u>. Vols. 1–3. Translated by William Hazlitt. New York: D. Appleton & Co.

    6. Harkness, J. , Van de Vijver, F. J. R., & Johnson, T. P. (2003b). Questionnaire design in comparative research. in J. A. Harkness, F. J. R. Van de Vijver, & P. Ph. Mohler<u>Cross-Cultural Survey Methods</u> (pp. 19-34). Hoboken, New Jersey: John Wiley and Sons, Inc.

    7. Harris, M. (1968). <u>The rise of anthropological theory</u>. New York: Thomas Y. Crowell Company.

    8. Hoodfar, H. (forthcoming, 2008). Family Law and Family Planning Policy in Pre- and Post-revolutionary Iran. Chapter 4 in K. M. Yount and H. Rashad (eds), <u>Family in the Middle East: Ideational Change in Egypt, Iran, and Tunisia</u>. Oxford, UK: Routledge.

    9. Kazuo, K. (1995). Born-again Chinese religion. <u>Asian Folklore Studies,</u> 54:127-130.

    10. King, G., Murray, C. J. L., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. <u>American Political Science Review, 98</u>(1), 191-207.

    11. King, G., & Wand, J. (2006). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. <u>Political Analysis Advance Access, 1093</u>(10), 1-21.

    12. Linshu, C. (1994). Studies on religions in modern China. <u>Numen,</u> 76-87.

13. Mandelbaum, M. (1971). <u>History, man, and reason: A study in nineteenth-century thought</u>. Baltimore: The John Hopkins Press.

14. Mill, James. (1848). <u>The History of British India</u>. Vol. 1. London: James Madden.

15. Millar, J. ([1779] 1979). The origin of the distinction of ranks. W. C. 1. 1. Lehmann (Ed.), <u>John Millar of Glasgow 1735-1801</u> . New York: Arno Press.

16. Nisbet, R. A. ([1969] 1975). <u>Social change and history</u>. New York: Oxford University Press.

17. Pennell, B.E., Harkness, J., Levenstein, R. & Quaglia, M. Challenges in Cross-National Data Collection. In <u>Survey Methods in Multinational, Multiregional and Multicultural Contexts,</u> edited by M. Braun, B. Edwards, J. Harkness, T. Johnson, L. Lyberg, P. Mohler, B.E. Pennell and T.W. Smith. Hoboken, N.J.: John Wiley & Sons, forthcoming.

18. Sadana, R., Tandon, A., Murray, C. J. L., Serdobova, I., Cao, Y., Xie, W., Ustun, B. & Chatterji, S. (2001). Describing population health in six domains: Comparable results from 66 household surveys.

19. Sanderson, S. K. (1990). <u>Social evolutionism. A critical history</u>. Oxford: Basil Blackwell.

20. Smith, A. D. (1973). <u>The concept of social change</u>. London: Routledge & Kegan Paul.

21. Smith, T. W. (2003). Developing comparable questions in cross-national surveys. in J. A. Harkness, F. J. R. Van de Vijver, & P. Mohler<u>Cross-Cultural Survey Methods</u> (pp. 69-91). Hoboken, New Jersey: John Wiley and Sons, Inc.

22. Stocking, G. W. Jr. (1968). <u>Race, culture, and evolution</u>. New York: The Free Press.

23. Stocking, G. W. Jr. (1987). <u>Victorian anthropology</u>. New York: The Free Press.

24. Tan, C.-B. (1983). Chinese religion in Malaysia: A general view. <u>Asian Folklore Studies, 42</u>, 217-25.

25. Teiser, S. (1995). Popular religion. <u>The Journal of Asian Studies,</u> 378-395.

26. Thornton, A. (2001). The developmental paradigm, reading history sideways, and family change. <u>Demography, 38</u>(4), 449-465.

27. Thornton, A. (2005). <u>Reading History Sideways: The Fallacy and Enduring Impact of the Developmental Paradigm on Family Life</u>. Chicago: University of

Chicago Press.

28. Tylor, E. B. (1871). Primitive culture. Vol. I). London: John Murray, Albermarle Street.

29. UNDP (United Nations Development Programme). (2003). Human Development Report 2003: Millenium Development Goals: A Compact Among Nations to End Human Poverty. New York: Oxford Press.

30. Yang, C. K. (1962). Religion in Chinese Soceity, A Study of Contemporary Social Functions of Religion and Some of their Historical Factors. American Sociological Review, 27, 439-440.

31. Yount, K.M. & Rashad, H. (eds). (forthcoming, 2008). Family in the Middle East: Ideational Change in Egypt, Iran, and Tunisia. Oxford, UK: Routledge.