Published in final edited form as:
Ageing Soc. 2011 January 1; 31(3): 355–370. doi:10.1017/S0144686X10000917.

The development of culturally-sensitive measures for research on ageing

BERIT INGERSOLL-DAYTON[*]

[*] School of Social Work, University of Michigan, Ann Arbor, USA

Abstract

Attempts to import existing measures developed in other countries when constructing research instruments for use with older people can result in several problems including inappropriate wording, unsuitable response sets, and insufficient attention to cultural nuances. This paper addresses such problems by discussing a mixed methods approach to measurement development (*i.e*. both qualitative and quantitative) that incorporates input from the aging adults for whom the measure is intended. To test this approach, a step-by step process to the development of a culturally-grounded measure for older Thai people is described. Using focus groups and in-depth interviews, the process begins with an identification of the culturally meaningful domains of the construct under study. Next, input is gathered from other studies; a preliminary quantitative measure is developed; the measure is reviewed by a panel of experts; and then it is pilot-tested. Cognitive interviews are utilized when pilot-testing of the items detects problems with measurement construction or interview methods. When these problems are remedied, the measure is incorporated into a large-scale survey and tested for its psychometric qualities. In addition to providing a template for culturally-sensitive measurement development in gerontology, this paper also highlights issues that researchers should consider when attempting to develop measures and provides suggestions for how to address such issues.

Keywords

Asia; measurement development; comparative research; culture; mixed methods

Introduction

Gerontology researchers around the world are increasingly drawn together in the face of global ageing. An area of common concern is the accurate measurement of older people's functioning and well-being. To address this concern, researchers may borrow standardised measures developed in other parts of the world. Despite careful attempts to translate and back-translate such measures into and from their own language, this approach results in several problems. One is that culturally-relevant dimensions of a construct that is under investigation may be overlooked if a measure is imported from another country (Christopher 1999; Ho 1998; Miller *et al*. 2006). A second problem is that the meanings of words may vary between cultures making it impossible to translate key concepts (Chaiyawat and Brown 2000; Kee 2007; Skevington, Bradshaw and Saxena 1999). A third is that the response sets for closed-ended items may not be suitable from one culture to another (Ingersoll-Dayton, Saengtienchai, Kespichayawattana and Aungsuroch 2004). Such problems are further

Address for correspondence: Berit Ingersoll-Dayton, School of Social Work, University of Michigan, Ann Arbor, MI 48109, USA bid@umich.edu.

compounded for gerontological researchers in non-Western cultures because they are often using measures developed for younger people in Western countries (*cf.* Krause 2002). To address these difficulties, researchers sometimes need to create their own measures. This paper describes the use of mixed methods in the construction of culturally-sensitive measures for aging populations and discusses some of the practical issues associated with measure development.

Using mixed methods (*i.e.* both qualitative approaches and quantitative approaches) has been gaining recognition in social science research because both have complementary strengths and weaknesses (Tashakkori and Teddlie 1998). Morgan (1998) provided a useful map for understanding the different ways in which qualitative and quantitative methods inform each other. As applied to measurement construction, he suggests that qualitative methods be initially employed to identify relevant domains and appropriate wording that are subsequently used in the construction of quantitative closed-ended items. By incorporating input from individuals for whom the measure is intended, this process is more likely to result in survey items that are culturally meaningful and valid.

This combination of methodological approaches is particularly relevant to the development of culturally-sensitive measures in gerontology because it provides a means by which to incorporate the perceptions of indigenous older adults. As yet, however, there have been few efforts in the gerontology literature to describe a step-by-step approach to measurement using this mixed methods process. One exception is an article by Krause (2002) in which he provided a comprehensive measurement development strategy for older people from diverse racial backgrounds in the United States. Using multiple qualitative methods to inform quantitative methods, he used a nine-step approach for constructing measures of religion appropriate for both Black and White older adults. These steps were: '(1) focus groups; (2) in-depth interviews; (3) input from quantitative studies; (4) developing preliminary quantitative measures; (5) review by expert panel; (6) cognitive interviews; (7) pilot study; (8) nationwide survey; and (9) psychometric testing' (p. S266). Though Krause's measurement construction strategy was developed to address cultural sensitivity in a Western country, this paper extends his approach by applying it cross-culturally to an Asian country.

Here, I describe a field test of Krause's multi-method approach to the construction of a culturally-grounded measure for older people in Thailand. Included is a step-by-step elucidation of the approach as well as an examination of the benefits and problems associated with this mixed-method approach. In addition, the paper discusses issues that gerontology researchers should consider when deciding whether and how to develop a measure that is grounded in the culture for which it is intended.

Steps in developing a culturally-sensitive measure

A team of gerontology researchers comprised of three Thai colleagues (Chanpen Saengtienchai, Jiraporn Kespichayawattana and Yupin Aungsuroch) and myself set out to assess psychological well-being among Thai elders. To do so, we first located a measure of psychological well-being developed by Ryff (1989*a*, 1989*b*), based on her content analysis of writings by several well-known Western theoreticians and clinicians. The measure has good psychometric properties (Ryff and Keyes 1995), is widely used in the United States, and includes items related to six domains of psychological well-being: self-acceptance, positive relations with others, autonomy, environmental mastery, purpose in life, and personal growth.

The Thai members of the research team first attempted to translate the measure from English, but found it impossible to find Thai words that captured the meaning of some items.

Western concepts from this measure of psychological well-being (such as environmental mastery) have no equivalent meaning in the Thai language. The inability to find appropriate words with equivalent meanings is not uncommon when translating a measure from one language to another (Chaiyawat and Brown 2001; Skevington, Bradshaw and Saxena 1999). This problem led our team to question, as have others (Christopher 1999), whether the domains of psychological well-being might be culturally variant. Our question resulted in an effort to identify the key dimensions of psychological well-being relevant to older Thai individuals and to construct a measure that was grounded on these culturally-specific dimensions. To do so, we used an adapted version of Krause's nine-step process of measurement development. This process and the adaptations are detailed in the next section.

Steps 1 and 2: Using focus groups and in-depth interviews

Initially, the research team sought to identify culturally-meaningful domains of psychological well-being. While Krause collected data first from focus groups and subsequently from in-depth interviews, to save time the research team collected data from both sources simultaneously. Focus groups and in-depth interviews are distinct but complementary qualitative methods. In focus groups, researchers gather information using open-ended questions posed to a group of people. One benefit of focus groups is that the interaction among group members may stimulate the development of ideas for discussion (Knodel 1995; Morgan 1997), but there are also drawbacks. For example, when conducting research on care-givers of people with dementia, Mahoney and colleagues (2005) found that focus groups were problematic for Chinese care-givers due to their discomfort with public disclosure in a community setting. From his experience of using focus groups with older informants in Asia, Knodel (1995) cautions that they are not appropriate for collecting detailed personal information. Thus for some research questions, collecting data using individualised in-depth interviews is more suitable because during one-on-one interviews, researchers can more easily follow up responses to open-ended questions and obtain more personal information. A benefit of both of these qualitative methods is that the direction of the inquiry can be shaped by the research participants which facilitates the collection of culturally-meaningful information. When used in conjunction, the specific benefits of the methods complement each other.

In this initial phase, the research team conducted 23 in-depth interviews and talked with 44 individuals from five focus groups. All were Thai elders aged 60 and over who lived in rural and urban areas around Bangkok. They were referred by staff from local health care centres and were selected based on variability in levels of education and income. The open-ended questions that we posed were based on two sources: Ryff's (1989*b*) questions that asked respondents to identify the characteristics of a mature, well-adjusted elder, and Ingersoll's (1985) ethnographic research on life quality in rural Thailand led us to ask about sources of enjoyment and hopes for the future. The questions were similar for both the focus groups and the in-depth interviews, although we included more probes for personal insights during the in-depth interviews. Simultaneously gathering data from both sources allowed for subsequent modifications to both sets of interview schedules as we gathered data from the two sources.

Data from these interviews were tape-recorded, transcribed and translated into English. Members of the research team read either the English or the Thai transcripts and independently developed a coding scheme. Together, team members reviewed each others' coding schemes, developed a comprehensive set of codes, and then augmented the scheme with topics suggested by previous research on well-being (Ingersoll 1985; Ryff 1989*b*). Using the resulting coding scheme, two researchers (one spoke English and the other was bilingual in Thai and English) coded the English transcripts and two native Thai speakers coded the Thai transcripts. The bilingual speaker was able to help the team develop coding

categories that would lend themselves to comparable Thai and English translations. The team members subsequently met as a group to review the coded data. Five dimensions of psychological well-being emerged from the data: harmony, interdependence, acceptance, respect, and enjoyment (Ingersoll-Dayton *et al.* 2001). This initial phase of data collection identified the key ingredients of psychological well-being experienced by the Thai research participants. Most of these components involved relationships with others (*e.g.* harmony, interdependence, respect) in marked contrast to the Western measure of psychological well-being developed by Ryff (1989*a*) that centred largely on intra-personal components (*e.g.* autonomy, personal growth, purpose in life).

Step 3: Gathering input from other studies

The next task was to determine whether there was sufficient information from the in-depth and focus group interview participants about each of the five dimensions of psychological well-being that had emerged. An analysis of the transcripts indicated that 'respect' was the only dimension on which there was sparse information. Fortunately, this dimension was the focus of a previous qualitative study (Ingersoll-Dayton and Saengtienchai 1999). Older respondents in this previous study described various facets of respect including: being asked for their advice by younger people; and observing that these people followed their advice. Based on such information, the research team developed items during this third step that specifically addressed these facets of respect and included the following: 'The young people in my family ask for my advice' and 'The young people in my family follow my advice'. Our approach differed from that of Krause who conducted a preliminary test of some of his close-ended items in a quantitative survey, in that we used information from another qualitative study to facilitate the development of some close-ended items; that is, the additional information about respect from the earlier qualitative study allowed the research team to elaborate on what the participants in the focus groups and in-depth interviews told us about this dimension of psychological well-being.

Step 4: Developing preliminary quantitative measures

The task during this step was to use the focus group and in-depth interview transcripts to develop culturally-grounded statements related to each of the five dimensions of well-being. The team incorporated words and phrases from the research participants' statements, to develop close-ended items. Special attention was paid to capturing the variety of issues raised by the participants. For example, with respect to the interpersonal dimensions of 'harmony' and 'interdependence', two issues were particularly salient. For one of these issues, the importance of one-on-one relationships with each family member, we developed an item: 'People in your extended family take care of you'. For the second issue, the importance of relationships among family members, we developed an item: 'Members of your extended family care about each other'.

Step 5: Reviewing of measures by expert panel

Having completed a draft of the potential close-ended items, the next task was to gather input from experts who were selected using two criteria: their knowledge about Thai older people and about assessing psychological well- being. Four scholars with expertise in one or both of these areas rated each of the preliminary psychological well-being items on a scale ('1' 'item has no problem' to '5' 'eliminate item entirely'). In addition, they provided their opinions on the adequacy of the five dimensions identified from the focus group and in-depth interviews. The four experts concurred that the five dimensions were adequate, suggested revised wordings for some of the items to increase their clarity, and suggested some additional items as indicators of the five dimensions. Using their feedback, the list of close-ended items was revised accordingly.

Step 6: Pilot-testing

Although Krause employed cognitive interviewing before the pilot testing, our research team originally decided to skip the cognitive interviews because of time constraints, which later enabled us to determine its relative importance when we later decided to include it (see Step 7 below). When conducting the pilot test, a purposive sampling strategy was implemented. Staff in urban and rural care centres in Bangkok and three nearby provinces identified people aged 60 or more years who represented variability with respect to health, education and socio-economic status. The resulting survey sample had 477 individuals (295 women and 182 men) who were interviewed by Thai graduate students in nursing.

This pilot test detected several problems with our measurement construction and interview methods. Many survey respondents had difficulty understanding the psychological well-being items. One problem was that items with negative terms (*e.g.* not happy) were confusing because the response choices also included negative options (*e.g.* disagree, not at all, never). Further, the response choices made participants uncomfortable because they involved options that necessitated disagreeing with their interviewers. Ultimately, it appeared that participants' desire to agree with the interviewer was affecting their choice of responses. In addition, the presence of others during the interview was problematic, especially since several items in the measure of psychological well-being were related to these individuals (e.g, 'In your extended family, people get along well together'). Analyses of the data confirmed the problematic nature of our items and interview methods in that there was little variability among the responses and they were highly skewed toward positive psychological well-being.

Step 7: Conducting cognitive interviews

These problematic reactions to the psychological well-being items during the pilot test led us to reconsider the initial decision to skip cognitive interviews. Instead, we decided to undertake this approach which involves asking research participants to 'think aloud' as they provide their interpretation of each item (Genest and Turk 1981; Schwartz 1999). In so doing, it is possible to determine which items are confusing and to ask for assistance from the research participants in revising problematic items. The Thai team members conducted cognitive interviews with 30 people (20 women and 10 men) aged 60 or more years from Bangkok and two nearby provinces. Most were fairly well-educated (*i.e.* had completed fourth grade). In recruiting these individuals, efforts were made to insure that they had the capacity to think about the clarity of the close-ended items from their own perspective as well as from the perspective of older Thais with less education. Staff and volunteers at the village health-care centres and 'elderly clubs' (*i.e.* senior centres) identified individuals who they believed were reflective and could explain their thought processes. Each participant was asked detailed questions about half of the items to avoid causing fatigue. Findings from the cognitive interviews were used to revise the items which were subsequently presented to other cognitive interview participants. This process led to a number of important insights.

First, it was possible to insure that the items were accurate indictors of the underlying dimensions of psychological well-being identified by the focus groups and in-depth interviews. During the cognitive interviews, team members asked questions that were intended to serve as validity checks on each item. For example, based on a focus group discussion about the fun that the participants gained from singing, the team developed an item related to enjoying singing as an indicator of the enjoyment dimension of psychological well-being. To obtain information about other aspects of enjoyment, additional questions were posed in subsequent interviews, such as: 'Is the statement clear? If not, how could it be improved? What are some other ways of enjoying life?' The responses during these later interviews led to an expanded version of the enjoyment item (*i.e.* 'You like to sing or listen

to music'). When this item was validated with other interviewees, it appeared to be a more accurate indicator of enjoyment.

Second, it was possible to determine if the items were generally understood. The participants were asked to think about the meaning of each item and, where necessary, to assist in changing the wording of items to make them more understandable and consistent with the intended meaning. One example was in relation to the wording of family relationships. To identify a broad term for family that would be inclusive of extended family members and would be easily understood by older Thais, a question was posed: 'When you think about your family members, who are you thinking about?' In addition, a number of terms for extended family were suggested to the participants and they were asked, 'Which is the better way of talking about family members?' Ultimately, the word for which we were searching emerged spontaneously as an interviewee was showing a picture of her family, which included her brother, sisters, nieces and nephews. A team member who was observing the interview noted the term that she used for extended family. This word, after further cognitive interview testing, became the term that was used to describe family members in the psychological well-being items.

Third, it was possible to develop a more appropriate response set for the items. During the cognitive interviews, participants were asked to answer the close-ended items using a number of different response sets and to describe the reasons for their choices. Ultimately, however, it was by listening to the way in which interviewees spontaneously responded to the close-ended items that the research team discovered a consistently useful response set. Rather than describing their reactions in terms of agreement and disagreement (the original response set), the cognitive interviewees responded spontaneously with words such as, 'that's true' or 'that's not true'. Subsequently, a response set that used different levels of truth ('not at all true' to 'very true') was developed that the participants could more readily answer. During these cognitive interviews, participants explained that response sets concerning agreement/disagreement evoked responses about how things ought to be in contrast to response sets concerning truth/non-truth which evoked responses about their actual situations. This insight persuaded us to discard the original response set and replace it with one that incorporated degrees of truth.

Finally, it was possible to revise our interviewing methods. The cognitive interviews provided an opportunity to explore the extent to which the presence of others might influence the participants' answers. This issue is seldom mentioned in the research literature but is particularly relevant to the conduct of research with older people in developing countries, such as Thailand, where many live with their families, especially their adult children (Chayovan and Knodel 1997). Questions posed to the cognitive interviewee were: 'Do you feel comfortable answering my questions when others are around?' 'When I ask you about your family and neighbours, does their presence influence your response?' Almost all of the interviewees acknowledged that the presence of others influenced their answers. Further, they offered a number of interviewing suggestions that included direct methods (*e.g.* asking family members permission to talk with the older person alone) and indirect methods (*e.g.* explaining that the interview needed to be conducted individually, beginning with the older person and conducting abbreviated interviews with any others still present at the end of the interview or moving to a more superficial area of the interview when others were present and then shifting back to more sensitive topics after they were gone).

Step 8: Conducting a national survey

The next task was to conduct a large-scale survey with multiple purposes. One was to discover additional methods of creating privacy when interviewing in the presence of others.

A second was to reduce the number of items in our measure, thereby making it more useful to other researchers. Another was to compare the structure of psychological well-being for older people in Thailand as compared to the United States (for a more detailed description of the findings from this survey *see* Ingersoll-Dayton *et al*. 2004).

After incorporating the changes in item wording suggested by the cognitive interviews, the revised measure of psychological well-being was included in a survey of older adults in Central Thailand. Using a multi-stage selection process, the research team conducted the interviews in Bangkok and in districts of three randomly-selected provinces. District Health Officers identified the sub-districts and health centres within them. In each sub-district, the research team randomly selected the villages as well as the potential participants to be interviewed. The resulting sample had 460 participants aged 60 or more years (283 women and 177 men).

The research team discussed with the interviewers, who were Thai graduate nursing students, the importance of privacy when asking about psychological well-being. Training on this topic included a description of the approaches suggested during the cognitive interviews along with instructions to identify additional privacy-enhancing strategies that they discovered during the course of their interviews. Strategies that emerged during the survey data collection included: suggesting to the elderly respondent that they find a 'quieter place' to talk and moving the location of the interview; lowering their voice and moving physically closer; and asking respondents who could read to point to their chosen response rather than speak it aloud.

To reduce the number of psychological well-being items, a confirmatory factor analysis was conducted. Two criteria were applied to the trimming process. The first was that items were retained for each of the five dimensions of well-being identified by focus group and in-depth interview participants. The second was that items had acceptable loadings on their respective dimensions. Using these criteria, the team reduced the number of items from 35 to 15 such that each dimension had three items. Structural equation modelling was conducted to identify the structure of the resulting 15-item measure of Thai psychological well-being and to compare it to that of Ryff's measure. The best-fit model of psychological well-being among older Thai people had two facets. One was inter-personal (*e.g.* harmony, interdependence, respect) and the other was intra-personal (*e.g.* acceptance, enjoyment). This structure was markedly different from the model identified by Ryff and Keyes (1995) from research on adults in the United States. While their measure of Western psychological well-being identified a single global factor that was predominantly intra-personal, the Thai measure resulted in two distinct factors. Taken together, these findings suggest that, like their Western counterparts, older Thai people experience an intra-personal component of psychological well-being, but unlike their Western counterparts, they also experience a well-developed inter-personal component of psychological well-being.

Step 9: Psychometric testing

The final task was to determine the reliability and validity of the newly-developed measure of psychological well-being. Using the data from the survey of older adults in Central Thailand, two indices were created – one for each facet of psychological well-being. The internal consistency for both subscales was adequate (alpha = 0.82 for inter-personal well-being; alpha = 0.69 for intra-personal well-being). A three-week test-retest reliability check was acceptable for the intra-personal well-being index (r = 0.68) but somewhat lower for the interpersonal well-being (r = 0.59). The validity of both subscales was acceptable in that they were positively correlated with a measure of life satisfaction (r= 0.31 for inter-personal well-being; r = 0.27 for intra-personal well-being) and negatively correlated with a measure of geriatric depression (r = -0.38 for inter-personal well-being; r = -0.39 for intra-personal

well-being). Similar to the Western measure of psychological well-being (Ryff and Keyes 1995), the distribution of both subscales was skewed toward positive psychological well-being. This skewness may be a function of the good psychological well-being of the tested sample, and it is possible that the inclusion of more vulnerable people, such as those who are hospitalised or in long-term care facilities, or the addition of more items are needed to produce a wider distribution of the measure's scores.

Issues in the development of culturally-sensitive measures

The field test of Krause's multi-modal approach to measurement development described here suggests that this process provides a useful template for constructing culturally-grounded measures. Our adaptations of his approach resulted in three important lessons for gerontology researchers. First, it may be preferable to conduct the first two steps at the same time (*i.e.* gathering focus group data and gathering in-depth interview data). In so doing, each method of data collection can assist in the modification of the interview questions for the other. Also, by collecting data from in-depth interviews and focus groups simultaneously, it is possible to reduce the amount of time spent on recruitment and data gathering. Second, the refinement of close-ended items (*i.e.* Step 3) can be improved in multiple ways. Krause did so by including the items in a survey for quantitative testing; we did so by including findings from another qualitative study to augment the existing items. Finally, our experience shows that cognitive interviewing is a crucial step in the measurement development process. This approach can provide gerontology researchers with important insights into not only the wording of survey items but also into culturally-specific methods by which to conduct survey interviews.

There are a number of issues that gerontology researchers should consider as they make decisions about developing their own measures. The next section highlights two areas of tension: (1) developing measures carefully but doing so within realistic time constraints, and (2) developing measures that are culturally sensitive while also allowing for cross-cultural comparisons. Some possible courses of action are suggested with the caveat that further systematic testing of these suggestions is needed.

Addressing time constraints

The multi-method approach described here requires the investment of considerable time and energy. Krause spent three years developing his measures of religiosity. My colleagues and I spent four years on the construction of the measure of psychological well-being described here. Our task may have taken longer as a result of alternating between data gathering for several weeks each year in Thailand and then meeting in the United States for a few weeks to analyse the data and plan the next step. It is possible to reduce the time devoted to measurement construction by limiting some of the activities associated with specific steps in the process. One possibility is to gather data from fewer individuals in the initial phases. For example, Lau *et al.* (2003) conducted focus groups with only seven individuals as part of a multi-modal approach to define the components of life quality among Chinese elders who had experienced a stroke.

Another possibility is to limit the time devoted to qualitative analysis. Researchers might, for example, selectively choose to transcribe and analyse only those in-depth interviews and focus-group interviews that provide the richest information. Alternatively, they may choose to modify an approach, such as cognitive interviewing. To illustrate, when constructing a culturally-grounded measure of psychiatric distress in a conflict zone in Afghanistan, Miller and his colleagues (2006) adapted the one-on-one approach commonly used in cognitive interviews. Instead, their research team met as a group to review and modify items in

conjunction with two Afghan consultants who could provide detailed feedback on items based on their knowledge of the indigenous population.

A third possibility is to forego tape recording and transcribing the focus group and in-depth interviews. This shortcut has been used when developing measures under very tight time constraints. For example, when constructing the Afghani measure of psychiatric distress, Miller and his colleagues (2006) relied on handwritten notes rather than tape recording and transcribing. Such shortcuts contributed toward the ability to complete their measure in 15 days. A drawback, of course, is that these different kinds of time-limiting strategies are less systematic and may result in useful information being overlooked. Another way to limit the time spent on measurement construction is by deleting one or more steps in the process. For example, researchers might decide to choose between conducting focus groups or in-depth interviews. If so, it is important to think about the benefits and limitations of each method. When trying to obtain data on sensitive issues which may not lend themselves to group discussion, in-depth interviews may result in more information (Knodel 1995; Krause 2002).

Enhancing the comparability of culturally-sensitive measures

Another issue to consider when developing culturally-sensitive measures is the extent to which such measures are culturally bounded. That is, when measures are developed with and for specific populations, the measures are valid for that specific population and should not be applied to other populations without further validation (Tran 2009). To illustrate, the measure of psychological well-being described in this paper may be bounded by the religious background of the Thai Buddhist elders involved in its development, and therefore it may not be applicable to Thai Muslim elders.

While the focus of this paper has been on developing measures that are sensitive to specific cultural groups, there are also ways in which to expand such measures to make them more cross-culturally relevant. One way is develop measures that include universal dimensions of a construct while also including dimensions that are culturally specific. For example The World Health Organization (WHOQOL Group 1994) has developed a quality of life instrument that is both culturally specific and culturally generalisable. To do so, they involved researchers from several countries who developed a core set of quality of life domains and items that were subsequently revised on the basis of focus groups conducted in the same countries. Next, focus groups from each country were used to identify additional country-specific quality of life indicators and a panel of experts was convened to determine what additional items should be added to the core set of items to maximise their cultural specificity for each country. In addition to these culturally-specific items, the WHOQOL researchers accounted for cultural variation in response sets as well. That is, they used standard anchor points (*e.g.* ‘very satisfied’ to ‘very dissatisfied’) and then chose relevant language to describe the intermediate points that were culturally specific to each country. This process avoided the problems associated with simply translating a response set from one language to another and resulted in a set of items that could be used for culturally specific as well as culturally comparative purposes (WHOQOL Group 1995, 1998).

Tran (2009) has recently described another process by which to develop measures that are both culturally sensitive and cross-culturally comparable. He suggests that the process begin with a meeting of experts from each of the cultures for which the measure is intended. This meeting of cross-cultural experts should result in an agreement on the key constructs of interest, a pool of potential items for the measure, and questions for the culturally-specific groups that will help refine these items. The pool of items and the questions are then reviewed within the culturally-specific groups by working through a number of steps (*e.g.* consensus among cultural-specific experts, focus groups, and cognitive interviews). After completing these cultural-specific steps, the findings are synthesised into items that are

further revised after pilot testing in each of the cultures for which the measure is intended. Finally, the resulting items are included in a large-scale survey for all the cultural groups to evaluate the measurement equivalence across the cultures. Taken together, these suggested steps point to another way of developing measures that are culturally specific and culturally comparative.

In a world that is rapidly ageing, it is vital that gerontology researchers have valid measures to identify areas of need and areas of progress. The measures must be sensitive to cultural nuances so that they can contribute to the enhancement of culturally-relevant programmes and services for older adults. This paper had described a process by which researchers can construct culturally-sensitive measures, and it has discussed some of the tensions involved in the measurement construction process. In particular, gerontological researchers need to be mindful of the multiple steps required by careful measurement development while still attending to their own realistic time constraints. In addition, they may need to develop measures that are both culturally sensitive as well as comparable to other cultural groups. While some alternatives for addressing these tensions are presented in this paper, further systematic testing of these methods is needed as we seek to improve strategies for measurement development within gerontology.

Acknowledgments

I would like to thank Kalyani Mehta who made this special issue possible; Chanpen Saengtienchai, Jiraporn Kespichayawattana, Yupin Aungsuroch, John Knodel, Neal Krause, Jasper Ingersoll and Fern Ingersoll who helped guide my thinking on this topic; and Ruth Dunkle and Deborah Keller-Cohen who provided useful comments on several drafts of this paper.

References

Chaiyawat W, Brown J. Psychometric properties of the Thai versions of state-trait anxiety inventory for children and child medical fear scale. Research in Nursing and Health. 2000; 23(5):406–14. [PubMed: 11052394]

Chayovan, N.; Knodel, J. A Report on the Survey of the Welfare of the Elderly in Thailand. Institute of Population Studies, Chulalongkorn University; Bangkok, Thailand: 1997.

Christopher JC. Situating psychological well-being: exploring the cultural roots of its theory and research. Journal of Counseling and Development. 1999; 77(2):141–52.

Genest, M.; Turk, DC. Think-aloud approaches to cognitive assessment.. In: Merluzzi, TV.; Glass, CR.; Genest, M., editors. Cognitive Assessment. Guilford Press; New York: 1981. p. 233-69.

Ho DY. Indigenous psychologies: Asian perspectives. Journal of Cross-Cultural Psychology. 1998; 29(1):88–103.

Ingersoll, J. Are they better off? Social well-being as a framework for judging development results; Paper presented at the Society for Economic Anthropology; Warrenton, Virginia. 1985.

Ingersoll-Dayton B, Saengtienchai C. Respect for the elderly in Asia: stability and change. International Journal of Aging and Human Development. 1999; 48(2):113–130. [PubMed: 10376957]

Ingersoll-Dayton B, Saengtienchai C, Kespichayawattana J, Aungsuroch Y. Psychological well-being Asian style: the perspective of Thai elders. Journal of Cross-Cultural Gerontology. 2001; 16(3): 283–302. [PubMed: 14617984]

Ingersoll-Dayton B, Saengtienchai C, Kespichayawattana J, Aungsuroch Y. Measuring psychological well-being: insights from Thai elders. The Gerontologist. 2004; 44(5):596–604. [PubMed: 15498835]

Kee, LH. Indigenising Social Work Practice in Sarawak. Strategic Information and Research Development Centre; Selangor, Malaysia: 2007.

Knodel J. Focus groups as a qualitative method for cross-cultural research in social gerontology. Journal of Cross-Cultural Gerontology. 1995; 10(1):7–20.

Krause N. A comprehensive strategy for developing closed-ended survey items for use in studies of older adults. Journal of Gerontology: Social Sciences. 2002; 57B(5):S263–74.

Lau. A, McKenna K, Chan C, Cummins R. Defining quality of life for Chinese elderly stroke survivors. Disability and Rehabilitation. 2003; 25(13):699–711. [PubMed: 12791555]

Mahoney D, Cloutterbuck J, Neary S, Zhan L. African American, Chinese, and Latino family caregivers' impressions of the onset and diagnosis of dementia: cross-cultural similarities and differences. The Gerontologist. 2005; 45(6):783–792. [PubMed: 16326660]

Miller KE, Omidian P, Quaraishy AS, Quraishy N, Nasiry MN, Nasiry S, Karyar N, Yaqubi AA. The Afghan symptom checklist: a culturally grounded approach to mental health assessment in a conflict zone. American Journal of Orthopsychiatry. 2006; 76(4):423–33. [PubMed: 17209710]

Morgan, DL. Focus Groups as Qualitative Research. Sage, Thousand Oaks; California: 1997. Not cited in text Now included in text on page 6

Morgan DL. Practical strategies for combining qualitative and quantitative methods: applications to health research. Qualitative Health Research. 1998; 8(3):362–76. [PubMed: 10558337]

Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. Journal of Personality and Social Psychology. 1989a; 57(6):1069–81.

Ryff CD. In the eye of the beholder: views of psychological well-being among middle-aged and older adults. Psychology and Aging. 1989b; 4(2):195–210. [PubMed: 2789747]

Ryff CD, Keyes CL. The structure of psychological well-being revisited. Journal of Personality and Social Psychology. 1995; 69(4):719–27. [PubMed: 7473027]

Schwartz N. Self-reports: how the questions shape the answers. American Psychologist. 1999; 54(2): 93–105.

Skevington SM, Bradshaw J, Saxena S. Selecting national items for the AHOQOL: conceptual and psychometric considerations. Social Science and Medicine. 1999; 48(4):473–87. [PubMed: 10075173]

Tashakkori, A.; Teddlie, C. Mixed Methodology: Combining Qualitative and Quantitative Approaches. Sage, Thousand Oaks; California: 1998.

WHOQOL Group. The development of the World Health Organization quality of life instrument (the WHOQOL).. In: Orley, J.; Kuyken, W., editors. Quality of Life Assessment: International Perspectives. Springer-Verlag; New York: 1994. p. 41-57.

WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organization. Social Science and Medicine. 1995; 41(10):1403–9. [PubMed: 8560308]

The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL) development and general psychometric properties. Social Science and Medicine. 1998; 46(12): 1569–85. [PubMed: 9672396]

Tran, TV. Developing Cross-Cultural Measurement. Oxford University Press; New York: 2009.