Research & Metrics
Population Services International
1120 Nineteenth Street, NW, Suite 600
Washington, D.C. 20036

Concept Paper

FoQus on Scales: Qualitative Methods for Improving Multi-Item Scale Development and Adaptation

© Population Services International, 2007

Contact Information

Hibist Astatke
Kate O'Connell
Steve Chapman
Research & Metrics
Population Services International
1120 19th Street, N.W. Suite 600
Washington, DC 20036
Tel: 202-785-0072
Fax: +1 202 785 0120
Email: hastatke@psi.org

Linda Kaljee
University of Maryland Baltimore
Department of Pediatrics
737 West Lombard Street
Baltimore, MD 21201
Tel: 410-706-0204
Fax: 410-706-0653
Email: lkaljee@peds.umaryland.edu

Table of Contents

Executive Summary **1**

Overview **3**

Theory, Multi-item Scales, and Qualitative Methods **4**

PSI PERForMance Framework and the PSI Behavior Change Framework *4*

Theory-based Interventions *5*

Use of Scales to Measure Constructs in PSI Behavior Change Framework *6*

Qualitative Methodologies for Scale Development and Adaptation *7*

Problem Statement **8**

How are multi-item scales currently developed for PSI studies? *8*

What is the problem with the current methods for developing scales? *9*

How does FoQus on Scales address the problem? *9*

An overview of FoQus on Scales *9*

A Methodological Framework for FoQus on Scales **11**

Adaptive and Emergent Approaches to Scale Development *11*

Qualitative Data Collection Tools for FoQus on Scales *15*

Data Interpretation and Item Generation Tools *17*

Implementation of FoQus on Scales Studies **19**

Country Level Implementation *19*

Advantages and Limitations *24*

Resources and Capacity *25*

Global Implementation *26*

Figures

Figure 1: PERForM and the PSI Behavior Change Framework 5
Figure 2: Adaptive and Emergent Approaches to Scale Development 14
Figure 3: **Option I** FoQus on Scales Implementation 21
Figure 4: **Option II** FoQus on Scales Implementation 22
Figure 5: **Option III** FoQus on Scales Implementation 23

Tables

Table 1: FoQus on Scales Implementation Resources 25

Executive Summary

FoQus on Scales is one component of Project FoQus (A Framework for Qualitative Research in Social Marketing). Project FoQus incorporates qualitative methods to improve different steps within the social marketing research process. FoQus on Scales is designed to improve the process for multi-item scale development and adaptation within the PSI Behavior Change Framework.

FoQus on Scales is designed to address the problem of minimal in-put from the target audience population in relation to scale development. Currently information useful for identifying and measuring relevant determinants is mostly being gathered from in-country programmers and researchers. This is problematic as scales developed without input from the target audience are less likely to be reliable and valid, limiting data analyses and interpretation and decreasing effectiveness of interventions programs. In addition, the current practice contradicts a major principal of social marketing which is to develop customer focused programming.

FoQus on Scales is developed to balance the 'etic' (outsiders) and 'emic' (insiders) perspectives throughout the process of scale development and adaptation. FoQus on Scales assumes that: 1) utilization of a conceptual framework increases the applicability of research findings and effectiveness of interventions and programs; 2) using multi-item scales to measure constructs increases validity and reliability of those measurements; and, 3) qualitative methods with significant in-put from the target audience contribute to the development and adaptation of valid and reliable scales.

FoQus on Scales includes both 'emergent' and 'adaptive' approaches. The 'emergent' approach exclusively includes the development of new scales to measure both existing and new constructs within the PSI Behavioral Change Framework. The 'adaptive' approach is based on the modification of existing scales with the option of scale development for new constructs which emerge during the qualitative data collection and analysis and/or where appropriate scales do not exist.

FoQus on Scales uses qualitative data collection methods including focus groups, individual one-on-one interviews, and written narratives as well as targeted activities such as free-listing, rankings, pile-sorts, and cognitive interviews to pre-test scales.

FoQus on Scales is implemented through options that allow for increasingly more input from members of the target population over time and gradually increase the amount of resources and technical capacity needed to implement FoQus on scales. This provides flexibility for in-country researchers and program managers in terms of available resources, experience, and project objectives.

Overview

Multi-item scales are useful for measuring concepts like attitudes and beliefs that can not be directly observed. They consist of a number of statements that allow for the measurement of complex concepts that are not easily measured by a single question. Currently 83% of all TRaC SMRS reports published in 2006 and 2007 report data collected using multi-item scales. The quality of these scales however varies greatly across programs and influences the usefulness of research findings for Evidence Based Decision Making.

A recent review of HIV/AIDS SMRS reports published in 2005 & 2006 revealed several issues that affect the quality of scales. These include conceptual issues such as items that don't reflect the PSI Behavior Change Framework, scales that do not delineate subscales, and scales that address more than one behavior (e.g. attitudes about condom use, partner reduction and abstinence). Technical issues identified include use of statements that include more than one idea and use of the name of determinants in statements within the scale.

PSI launched its initiative to start using multi-item scales in the summer of 2004. PSI uses multi-item scales to measure determinants of behavior in the PSI Behavior Change Framework (see concept paper on the PSI Behavior Change Framework). This framework guides research that can be used to design and evaluate social marketing programs, as well as moving toward the generation of theory and theory-based programming.

FoQus on Scales is an initiative designed to improve the process for multi-item scale development and adaptation within the PSI Behavior Change Framework using qualitative methods. In this paper, we will first present the PSI Behavior Change Framework followed by a review of the justification for theory-based studies to design and evaluate interventions, utilization of scales to measure complex constructs within a theoretical framework, and use of qualitative methodologies as a component in scale and construct adaptation and development. We will then outline a methodological approach and procedures for FoQus on Scales including alternative methodological options for the implementation of FoQus on Scales.

Theory, Multi-item Scales, and Qualitative Methods

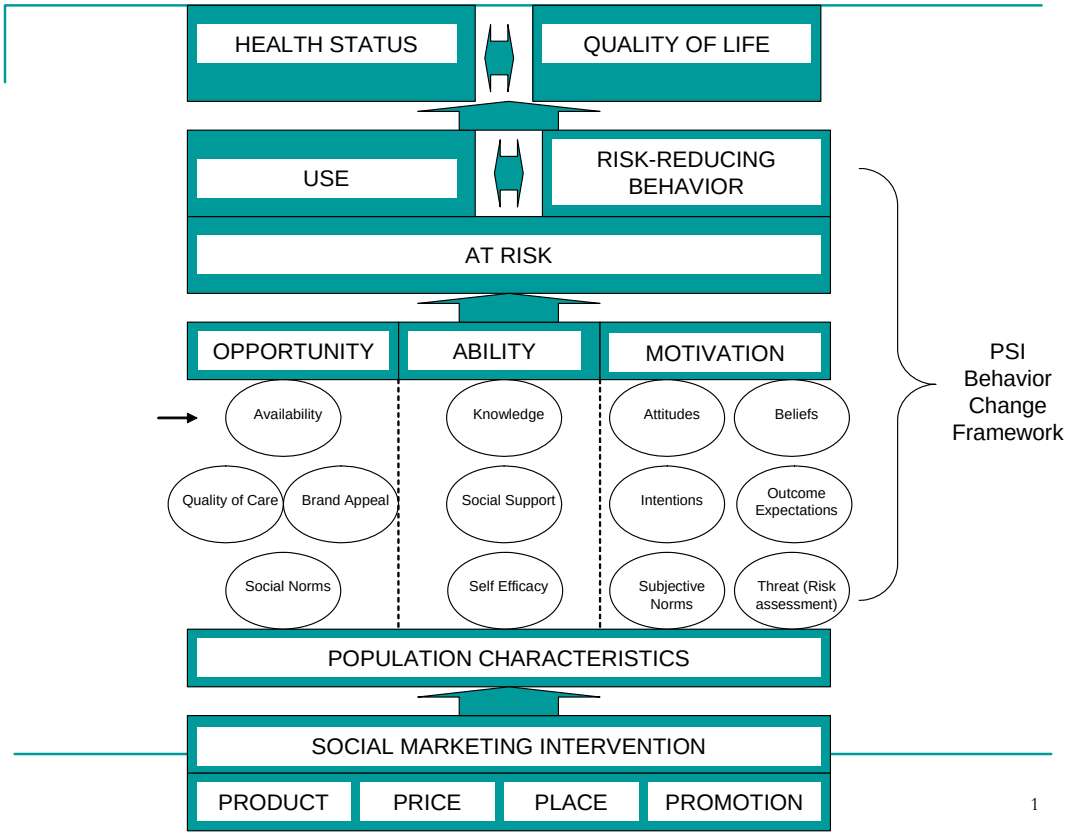
PSI PERForMance Framework and the PSI Behavior Change Framework

PSI research and programming is guided by PERForM (a Performance Framework for Social Marketing). PERForM defines measures of social marketing performance (Chapman, 2004). The PSI Behavior Change Framework is embedded within PERForM (See Figure One). The relationship proposed by the PSI Behavior Change Framework guide segmentation and evaluation studies as well as the development and validation of multi-item scales in PSI TRaC studies (Chapman & Coombes, 2003; Patel & Chapman, 2005).

The PSI Behavior Change Framework defines **opportunity**, **ability** and **motivation (OAM)** and situates determinants within these 'summary constructs' in relation to the use of health products or risk reducing behaviors (Patel & Chapman, 2004). Opportunity is defined as institutional or structural factors that influence an individual's chance to perform a promoted behavior. Ability is an individual's skills or proficiencies needed to perform a promoted behavior. Motivation is an individual's arousal or desire to perform a promoted behavior.

Sixteen behavioral determinants ("bubbles") deemed to be relevant for a wide range of health behaviors are included in the PSI Behavior Change Framework based on research and theory from several disciplines including the fields of psychology, health communication and social marketing (Patel & Chapman, 2004). After initial development of the PSI Framework, expert opinions were obtained to help with revisions and modification to the framework. Determinants within the PSI Framework were categorized within the OAM summary constructs. Determinants were only included if they were 'mutable' (able to be changed) and measurable.

Figure 1: PERForM and the PSI Behavior Change Framework



Theory-based Interventions

PSI's efforts to develop and implement a framework that could be used to design Social Marketing interventions began in 1999 (Patel & Chapman, 2004). This decision was in part based on a need to develop more effective and efficient interventions. At this time the PSI Behavior Change Framework provides a model for organizing health-related behavioral determinants within the OAM summary constructs. Through the implementation of FoQus on Scales, PSI will move further toward behavior-change theory development and testing.

For more than 15 years, there has been an increasingly strong literature-base to support the need for theory-based program interventions for a wide range of health issues including engagement in exercise and nutrition, smoking cessation, diarrheal disease, and HIV/AIDS (Lee, Avis, & Arthur 2007; Cole, Waldrop, D'Auria, & Garner, 2006; Simmons & Brandon 2007; Stanton, Black, Engle, & Pelto, 1992; Fishbein, 2000.)

Over this time period, evidence has accumulated that theory-based interventions are more efficacious than those interventions developed outside of behavioral and behavioral change theories. This evidence includes both U.S. based and international health education and promotion activities, and includes studies across a wide range of population groups (Kim, Stanton, Li, Dickersin, & Galbraith 1997; Graves, 2003; van Empelen et al. 2003).

These studies indicate several advantages of using behavior change theories when designing and evaluating prevention programs. First, theory based research tools and interventions increase the generalizability and replicability of the research and the interventions (Bradley, Wiles, Kinmonth, Mant, & Gantley 1999) This allows for valid and reliable cross-population comparisons, and decreases the need to recreate assessment and evaluation tools, and interventions for each unique audience. Second, the inclusion of psychosocial measures which are linked to theoretical constructs provide the means to more accurately describe changes at post-intervention, which again increases replicability of an intervention program (van Empelen et al. 2003). Third, the use of a theory increases researchers' chances of developing valid measures (DeVellis 2003). Fourth, theory-based research improves the ability to generate and test hypotheses regarding behavioral change.

Use of Scales to Measure Constructs in PSI Behavior Change Framework

Multi-items scales are used to measure the determinants within the PSI Behavior Change Framework because they are likely to be more reliable than single items and can be checked for their reliability and validity through statistical tests. **Reliability** is the extent to which an instrument will produce the same results when used two or more times and reflects the accuracy of the scale. **Validity** is the extent to which an instrument measures what you would like it to measure.

Multi-item scales are likely to be more reliable over time than single items because of the scope and precision of information gathered (Spector, 1992). The variety of statements used to get at different aspects of a concept increases the **scope.** The fact that scales have different response options and use several items to measure a concept allows for a larger variation in terms of the choices that respondents have and thus more **precision** in responses.

Qualitative Methodologies for Scale Development and Adaptation

Scales are often developed for use within a particular population, most frequently in Western industrialized countries. These scales are often limited in terms of their validity and reliability in other populations, and may lack items or sub-scales, which are salient to populations from developing countries, or from other cultural groups (Camfield & Ruta, 2007; Tsai, 2002). Only through the use of qualitative methods can culturally sensitive and valid constructs and scales be developed from existing theoretical frameworks and measures (Chwalow, A. J., 1995). Qualitative methods also contribute to the translation process and ensure that items retain their original intent and are in written in the colloquial language of the target audience.

Qualitative methods are an integral part of multi-phase research for both development of new scales and adaptation of existing scales. These methods are used to elicit information from the target population on a specific topic (e.g. social norms for number of sexual partners) (George, Mackinnon, Kong, & Stewart, 2006; Chen, Chen, Lotus Shyu, & Tang, 2007). At level one, qualitative data provide statements which can be used to generate scale items. At a second level, these data can be used for adaptation and creation of determinants and modification/generation of behavioral theory (Greenhaigh, Chowdhury, & Wood, 2006).

Problem Statement

How are multi-item scales currently developed for PSI studies?

Currently PSI has two methods for developing multi-item scales for use in TRaC surveys. The first can be described as a "top bottom" approach where a Washington DC based or regional researcher develops a model questionnaire based on a literature review. This model questionnaire is then adapted to local country contexts by research managers and programmers. This usually takes place in a workshop or meeting where researchers and programmers propose ways to make scales more locally relevant by adding and/or dropping items and changing the wording or phrasing of statements used. Currently model questionnaires have been developed for the following behaviors in the areas of HIV/AIDS, malaria and family planning: condom use, injection drug use, contraceptive use and ITN use. A variation of this method involves adaptation of a questionnaire that has been developed for use in one country for use in another country.

The second method can be described as a "bottom up" approach for development of multi-item scales. With this approach in country programmers and researcher, often with guidance from a regional or Washington DC based researcher, develop items in a workshop style meeting. Both methods utilize the PSI Behavior Change Framework and allow for identification of which determinants within the Framework are relevant for the particular target audience and health issue, and the modification of item(s) used to measure those determinants.

These methods reflect varying degrees of a combination of an "etic" and "emic" approach to scale development. An "etic" approach assumes that the same concepts are relevant across different contexts and that they can be defined and measured the same way across sites. An "emic" approach assumes little overlap across contexts and cultures and focuses on the adaptation and/or development of specific concepts and measures through data collected from members of the target audience (Berry, J. W. & Dansen, 1974).

What is the problem with the current methods for developing scales?

Within both methods programmer and researcher knowledge and experience (a primarily "etic" perspective) has been the main source of information for adapting and developing multi-item scales. In only a small number of cases qualitative data from studies conducted with the target audience have been used to develop and/or adapt multi-item scales. Thus, PSI's approach to scale development often does not include or includes only minimal input from members of the target audience ("emic" perspective). This is problematic as scales developed without input from the target audience are less likely to be reliable and valid, which limits data analyses and interpretation and decreases effectiveness of interventions programs. In addition, the current approach contradicts a major principal of social marketing which is to develop customer focused programming.

How does FoQus on Scales address the problem?

FoQus on Scales provides a methodological framework for PSI programs to increase the "emic" perspective in the processes of theory generation, defining determinants, and scale development and adaptation. Through qualitative research, FoQus on Scales improves the quality of scales developed for use in TRaC studies. Methods such as focus group discussions and one-on-one in-depth interviews are used to gather information from members of the target audience and: 1) gauge the relevance of the determinant in the PSI Behavior Change Framework for the particular health issue and population; 2) determine and describe additional determinants not included in the PSI Behavior Change Framework;; 3) describe ways in which the determinants are expressed and can be measured in local contexts; and, 4) modify the existing PSI Behavior Change Framework with the objective of theory generation and testing. These data are used to develop reliable and valid multi-item scales that are reflective of contextual issues.

An overview of FoQus on Scales

FoQus on Scales is one component of Project FoQus (A Framework for Qualitative Research in Social Marketing). Project FoQus incorporates qualitative methods to improve different steps within the social marketing research process. The other three components of project FoQus involve use of qualitative data to design interventions, develop concepts for intervention activities, and test these concepts. FoQus on Scales is a qualitative component used to develop

multi-item scales for use in the TraC (Tracking Results Continuously) study design. TRaC studies are used to design, monitor and evaluate social marketing programs.

FoQus on Scales uses multiple qualitative methodologies, which can be combined at increasing levels of complexity in order to assist in the development and/or adaptation of determinants and scales. This phased approach allows for an initial introduction of methods that require lower technical capacity and less resources (see Section on Implementation of FoQus on Scales Study). Scale development and adaptation includes: 1) modifying the PSI Behavior Change Framework including defining determinants to reflect local contexts; 2) generating or modifying items to be included in the scales to accurately measure these determinants; 3) creating meaningful response options for the target population; and, 4) developing tools to aid respondents, e.g., boards with photographs to help respondents understand Likert scale options..

A Methodological Framework for FoQus on Scales

FoQus on Scales combines 'etic' and 'emic' data, and thereby multiple perspectives and interpretations into theory-based determinants and scales. The 'etic' perspective is based in the application of the PSI Behavioral Change Framework and data interpretation from PSI staff. The 'emic' perspectives are drawn from qualitative interviews with members of the target audience and workshops with local in-country researchers and programmers.

Adaptive and Emergent Approaches to Scale Development

FoQus on Scales will be implemented in one of two ways: 1. an '**emergent**' scale-development approach; and, 2) an '**adaptive**' scale-modification approach. Each approach includes three phases of preparation, data collection and analysis, and application (See Figure 2).

Qualitative theory and methods are based on the argument of the need to start research as an inductive exercise whereby through an iterative process of data collection, discovery, and analysis models and theoretical frameworks are developed. This approach begins with open-ended qualitative data collection with limited constraints from the researcher's perspective (the 'etic" perspective) and theoretical models (Miller & Fredericks, 1999).

The FoQus on Scales 'emergent' approach is a modification of this qualitative inductive research model. The 'emergent' approach starts with a literature review and/or discussions with local in-country program and research staff regarding behavioral determinants and relationships between determinants and behaviors. This exploration phase is focused on determinants, which are salient to the particular health issue and target audience. The emergent approach is of particular value in studies where little or no existing data are available regarding a particular target audience and health issue. Pre-existing scales are not used in the 'emergent' approach, but rather through the qualitative data collection and analysis, determinants are identified and scales are developed.

Qualitative data collected from the target audience are used to modify the PSI Behavior Change Framework, define new determinants and/or redefine existing determinants, and develop new scales for measuring these determinants. These determinants and scales are therefore specific to the culture of the target audience and for the specific health issue. The 'emergent' approach provides a 'bottom up' approach to scale development for context specific determinants.

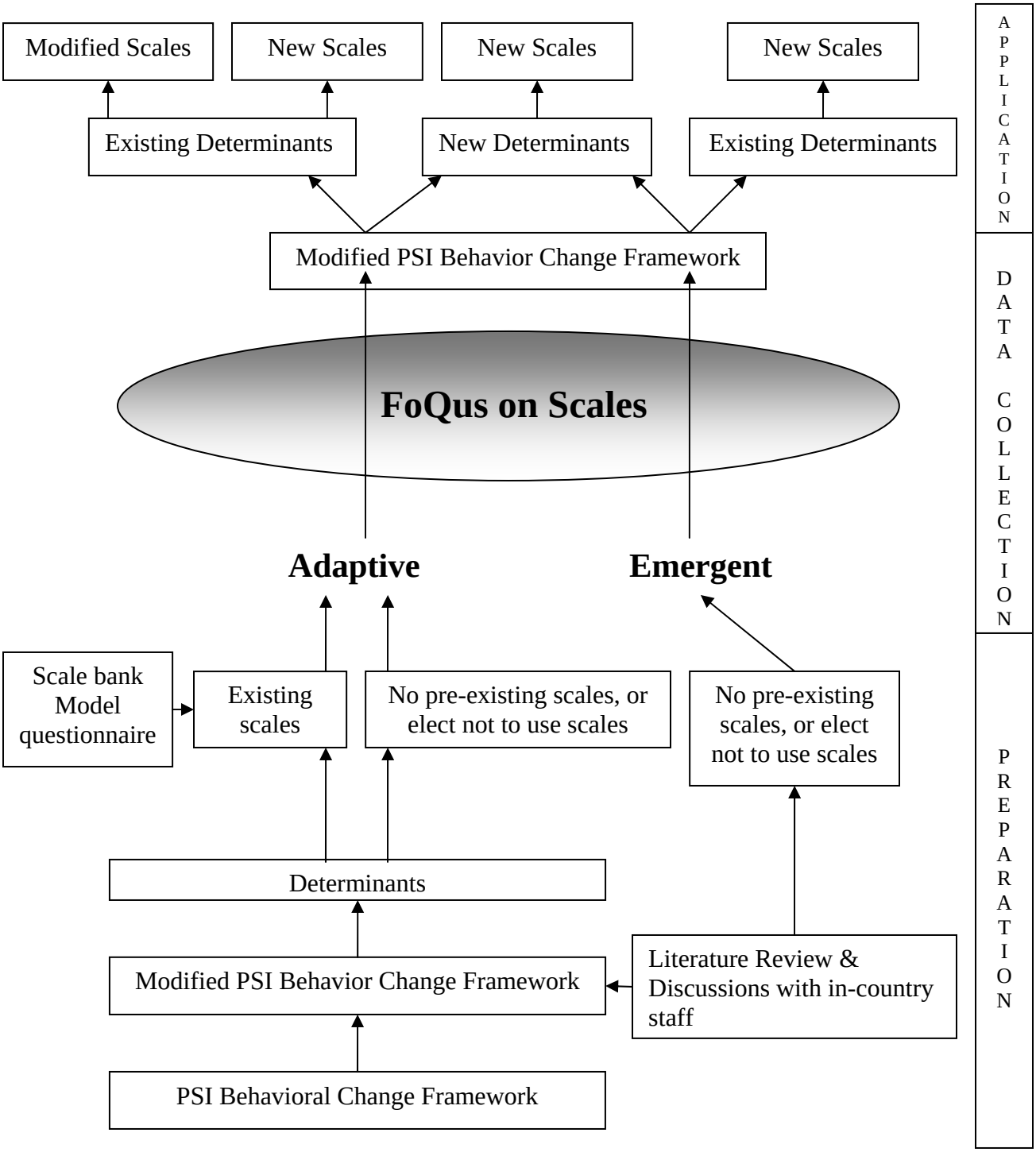
The ‘adaptive’ approach begins with the PSI Behavioral Change Framework and a literature review and/or discussions with local staff. In the ‘adaptive’ approach, the PSI Behavioral Change Framework is adapted for the specific health issue and target audience, and determinants are chosen to include prior to the qualitative data collection phase. Unlike the ‘emergent’ approach which creates new scales for both existing and emerging determinants, the ‘adaptive’ approach begins with existing determinants or existing scales from model questionnaire and/or the PSI scale bank. The PSI scale bank includes scales from a comprehensive literature review and PSI HIV studies from the Social Marketing Research Series reports (2005-2006). Most of these scales have been tested in one or more populations, but require culturally specific adaptations to increase validity and reliability. The same qualitative data collection and analysis methods are employed as for the ‘emergent’ approach, except the focus is on modifying existing determinants. However, the ‘adaptive’ approach does not rule out the development of new determinants within the PSI Framework, or the development of new scales.

The ‘emergent’ approach is more labor intensive and requires more expertise in scale development than the ‘adaptive’ approach. The ‘adaptive’ approach recognizes the potential time and resource-saving benefits of using existing questionnaires and scales, but also the need for modification based on culturally specific perspectives and attitudes.

Both the ‘emergent’ and the ‘adaptive’ approaches differ from the existing PSI model for scale development. Currently most scales are developed with in-put from program staff and researchers but not members of the target audience. Both approaches within FoQus on Scales depend on data from the target population as the primary source for the adaptation and development processes increasing the likelihood that determinants and scales reflect the ‘emic’ perspective. This increases the strength of the association between determinants and scale items, and thereby increasing validity and reliability of scales. Utilization of either the FoQus on Scales ‘emergent’ or ‘adaptive’ approach within the TraC study design will improve the quality of multi-items scales and provide the means toward more sophisticated quantitative data analysis including analysis that will allow for examination of causal relationships between determinants and behaviors in the PSI Behavior Change Framework. This should allow researchers to examine the validity of multi-items scales. Use of these approaches should also provide data that can be used to test hypothesized relationships between determinants and generate behavior change theory. These data on the mechanism for determinants mutable by social marketing interventions will be

useful for developing more effective and efficient programs and will improve segmentation studies, intervention evaluations, and social marketing strategies.

Figure 2: Adaptive and Emergent Approaches to Scale Development



Qualitative Data Collection Tools for FoQus on Scales

Qualitative data methodologies are based on assumption that it is important to capture the perspectives of a particular population as presented in the language of that population. The underlying assumption behind qualitative methods is that it is important to understand attitudes, perspectives, and behaviors within a social context. Qualitative data collection and analysis is an iterative process, allowing for changes in questions asked, interpretations of findings, and theoretical frameworks.

Focus group or one-on-one interview data will be used to develop and/or adapt determinants within the PSI Behavior Change Framework, and generate/assemble or adapt scale items to measure these determinants. Qualitative methodological tools for scale development include 'free listing' and 'pile-sorting', which will be designed to elicit data focused on determinants from the audience population and their interpretations of relationships among these determinants and between determinants and behaviors.

Qualitative Data Collection Methods

Focus groups have several advantages for collection of applied qualitative data. First, data collection is cost and time effective and the focus group interview increases sample size. Second, the interactions among participants can enhance the data and provide information on how individuals interact within the social setting. Focus groups are a way of revealing social processes and the ways in which these processes are collectively shaped. Third, individuals often enjoy participating in the group discussion and feel more comfortable than in one-on-one interviews (Patton 2002).

One-on-one semi-structured interviews are an important method for obtaining more personal information about respondents particularly on sensitive topics. Individual interviews also allow for alternative views on a subject, as focus groups respondents frequently express majority opinions.

Both focus group and one-on-one interviews use **semi-structured interview guides** to provide an outline of topics and questions to be included during the discussion. These interviews are conducted in a 'conversational style', and questions are not always asked in the same order, nor

necessarily worded in the same way. The semi-structured interview guides should be seen as dynamic, as new questions may emerge during data collection, or questions may be found to be repetitive or not relevant to the population.

Structured Interviewing Techniques

Free listing is a process for facilitating collection of data for construct and scale development and adaptation. Free Listing will be used to ask focus group participants to list beliefs, attitudes, perceptions within a certain topic area, e.g., reasons why adolescents don't use condoms. Participants will be asked additional questions to elicit further explanations and contexts for the listed statements. Free-listing will assist in the generation and elucidation of statements from the target audience for use in defining and identifying determinants, as well as for development of scale items.

Pile sorting is another qualitative tool that will be used. Pile sorting involves asking participants to take a set of statements and to sort them by some common denominator. Pile sorts will generally be used in one-one-one interviews. If sufficient time and resources are not available pile sorts may be used in group interviews. In the latter case one moderator provides instructions to three or four participants who each work individually on the pile sorting exercise simultaneously. Further data collection during pile sorting will include use of semi-structured interview guides to collect data on contexts for the categorizations, e.g., why certain items are together and how the respondents would name that category. Through pile sorting, statements generated from free-listing are categorized into determinants and scales.

Written narratives provide an opportunity for members of the target population to respond to specific questions in an alternative format to oral interviews. This methodology will be restricted to populations with a high level of literacy, and is most relevant when the topics are highly sensitive within the culture, e.g., pre-marital sexual relations. The narratives provide respondents with more time to reflect on complex issues and difficult questions. Written narratives will be used as follow ups to focus group discussions or one on one interviews since they are most effectively used in conjunction with other qualitative methods to clarify previously confusing or conflicting responses, to introduce and expand on emergent themes, and/or to provide a means to obtain greater detail about specific concepts.

Cognitive Interviews are used to pre-test multi-item scales. Group cognitive interviews are targeted 'focus groups' designed to elicit information about the understandability of scale items and appropriateness of response options for the population (Willis, 1999). Group interviews utilize a draft of the scale items, structured guides with questions on clarity of these items, assessment of agreement of intent of items between respondents and facilitators, language appropriateness and level of difficulty, and social desirability bias. Individual cognitive interviews are one-on-one interviews used to get detailed information about participants thought processes as they respond to multi-item scales. Individual cognitive interviews get similar information to the group interviews by asking participants to talk out loud as they respond to multi-item scales.

Data Interpretation and Item Generation Tools

Data interpretation and scale development workshops will be used to get input from in-country program staff as well as members of the target audience. These workshops will be lead by research mangers or consultants with technical assistance from a regional or Washington D.C. based researcher.

Data interpretation workshops will be used to involve in-country program staff and members of the target audience in the interpretation of the qualitative findings. At the beginning of the workshop participants will be provided with background information on the purpose of the study, data collection methods, qualitative data management and analysis as well as their role in the analysis process. Participants will work with data that has already been categorized into determinants of behavior. The participants will be asked whether or not they agree with the categorization and their reasons. . By the end of the session participants may suggest categorization of data under different determinants or they may create entirely new determinants.

Scale development workshops will also begin with an orientation to the scale development process and advantages to use of multi-item scales. Participants will be provided with opportunities to learn how to generate or modify statements to be used in multi-item scales. When data are being used to modify existing scales, participants will work on first revising the meaning and purpose of statements, adding or dropping statements, determining the number of dimensions or subscales and improving phrasing and word choice. Once statements have been revised the determinant definitions will be updated to reflect changes to the statements.

When data are being used to develop statements, participants will first define new determinants or update definitions for determinants in the PSI Behavior Change Framework and make decisions as to whether a one or multi-dimensional scale will be generated. Participants will then generate as many statements as possible for each dimension in a scale using its definition and qualitative data. These statements will be evaluated using criteria for writing “good” statements for multi-item scales. Once all statements have been finalized determinant definitions will be updated to reflect the content of the scales.

Implementation of FoQus on Scales Studies

A phased approach including **three** options for implementation of FoQus on Scales will be introduced across and within PSI country programs. This approach allows country programs to tailor FoQus on Scales to their needs and resources as well as to different research objectives and application. This approach also allows as many countries as possible to benefit from FoQus on Scales within the first year after its launch at the annual 2007 research retreat.

Country Level Implementation

FoQus on Scales is divided into three options that require increasing levels of technical capacity, resources, and input from the target audience. The objective of Option I is to generate or modify multi-items scales based on input from the target audience. An adaptive or an emergent approach of scale development is possible under this option. The objective of Option II is to generate multi-item scales and determinants that are defined by the target audience. An emergent scale development approach is recommended for this option, as an adaptive approach will underutilize the data collected with Option II. The objective of Option III is to utilize qualitative data to expand on either Option I or Option II, and further develop, modify, and/or revise scales after an initial quantitative pilot. The expanded component of Option III will use an adaptive approach.

Option I uses focus group discussions with free listing to identify relevant determinants of behavior and open-ended questions to gather contextual data for scale development/modification. Open ended questions are also used to explore the relevance of determinants identified through the PSI Behavior Change Framework, literature reviews and/or program and research staff. (see Figure 3). Under ideal conditions this process includes at least two sets of qualitative studies. The first set of studies is used to identify determinants of behavior through free listing and gather contextual information using open ended questions. Focus groups, one-on-one interviews, or narrative data can be used to generate contextual data. The second set of studies is used to explore the relevance previously identified determinants. Because of lack of resources and/or time, some country programs may elect to only conduct one set of focus groups and either identify determinants or explore the relevance of previously identified determinants but not both. Research staff will be responsible for analyzing data generated through Option I. Data analysis will include coding of the qualitative data, compiling these data, and writing summary statements. These data will be interpreted and scale items generated or adapted by PSI research staff only, or PSI research staff, in-country program staff and members of the target audience.

Option II begins with focus group discussions including free listing to identify relevant determinants of behavior and elicit phrases and statements for scale development and modification (see Figure 4). These statements will then be used in pile-sorting activities in one-on-one or group interviews. The target audience will categorize statements into piles (determinants) relevant for a particular behavior (e.g. condom use) and name and define these categories. Data from the pile sorting will be analyzed through aggregate proximity matrices using a software program (e.g. Anthropak), which provide a visual representation of 'clusters' of statements categorized by participants. Quantitative analysis of these data can further assist in determining significant differences or similarities between sub-populations, e.g., males and females in their clustering of the items (Borgatti, 2002). Through this exercise, statements generated from free-listing will be categorized into determinants and used by research staff to construct scales.

Option III begins with either Option I or Option II, and continues the use of qualitative data methodologies for the revision, adaptation, and/or creation of scales and determinants after an initial quantitative pilot and factor analysis. In Option III, there are two possible sources of data for scale refinement: 1) analysis of qualitative data from focus groups, one-on-one interviews, or written narratives collected during Option I or II; or 2) collection and analysis of new qualitative data (focus groups, one-on-one interviews, and/or written narratives) specifically targeted to assist with the interpretation of findings from the exploratory and factor analysis conducted with the quantitative pilot data.

In all three options multi-items scales are reviewed by an expert panel before being used in TRaC studies. Group or individual cognitive interviews are recommended but not required for pre-testing scales prior to the quantitative pilot. These cognitive interview data are used to make additional revisions to the scale items.

Figure 3: Option I FoQus on Scales Implementation

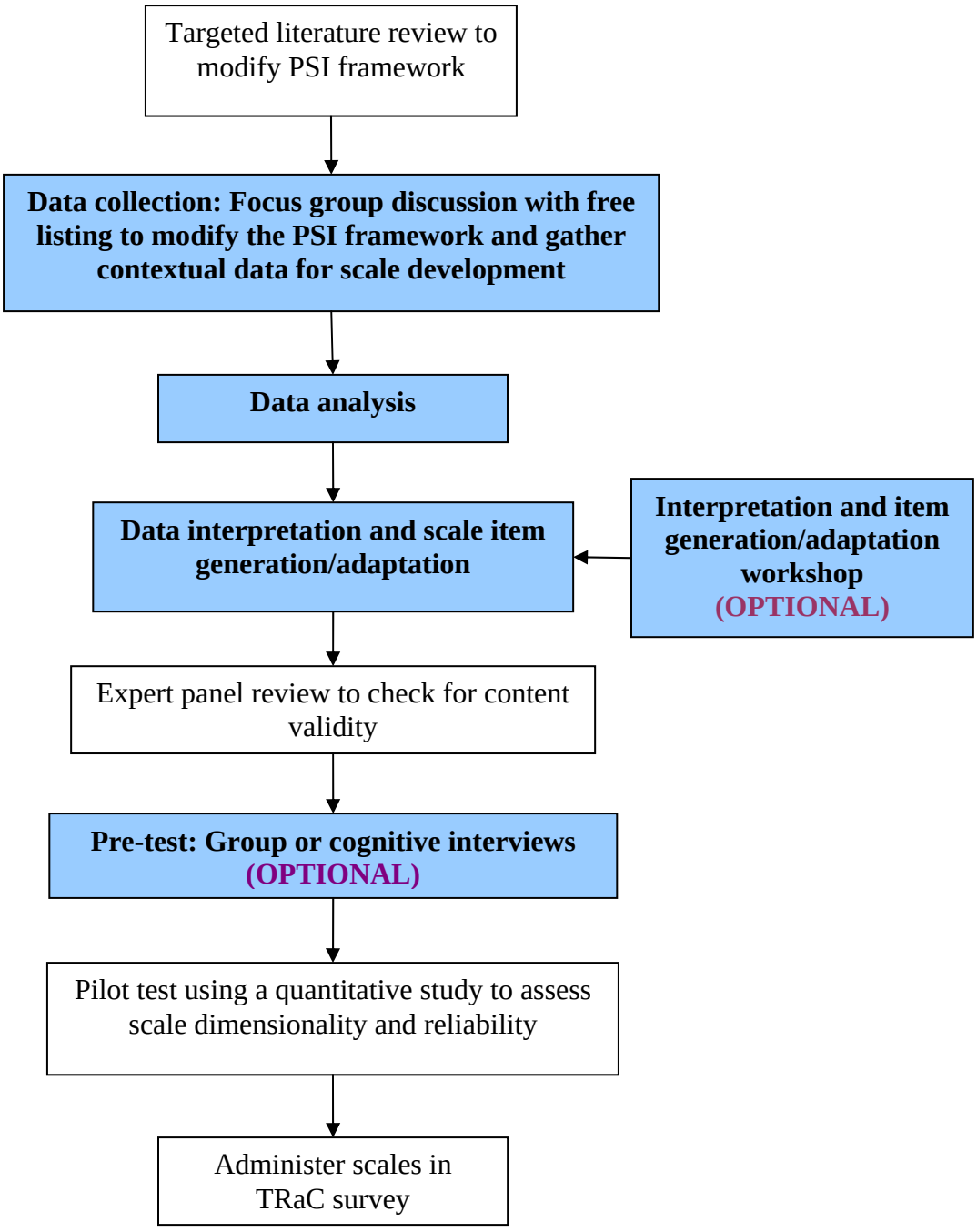


Figure 4 : Option II FoQus on Scales Implementation

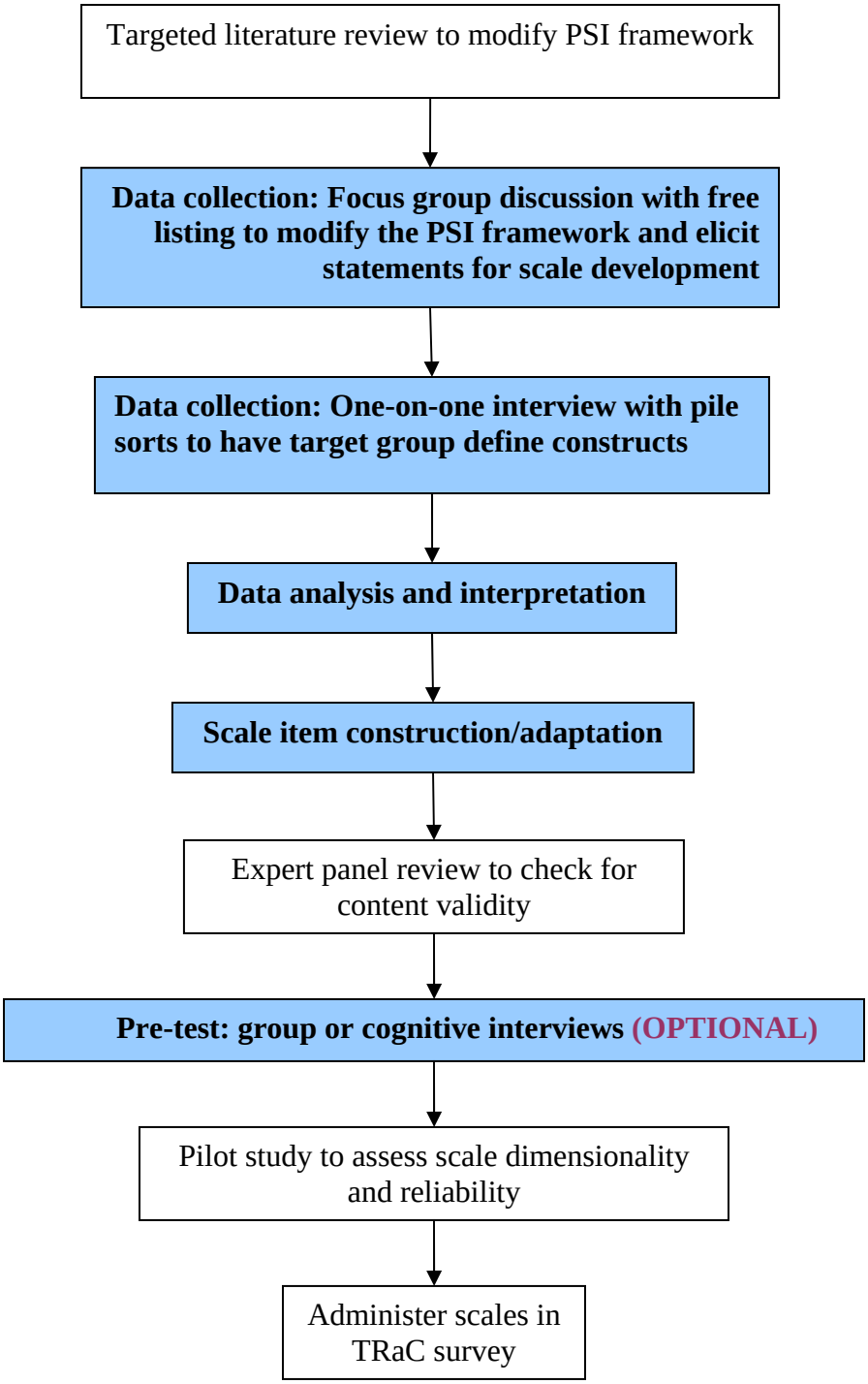
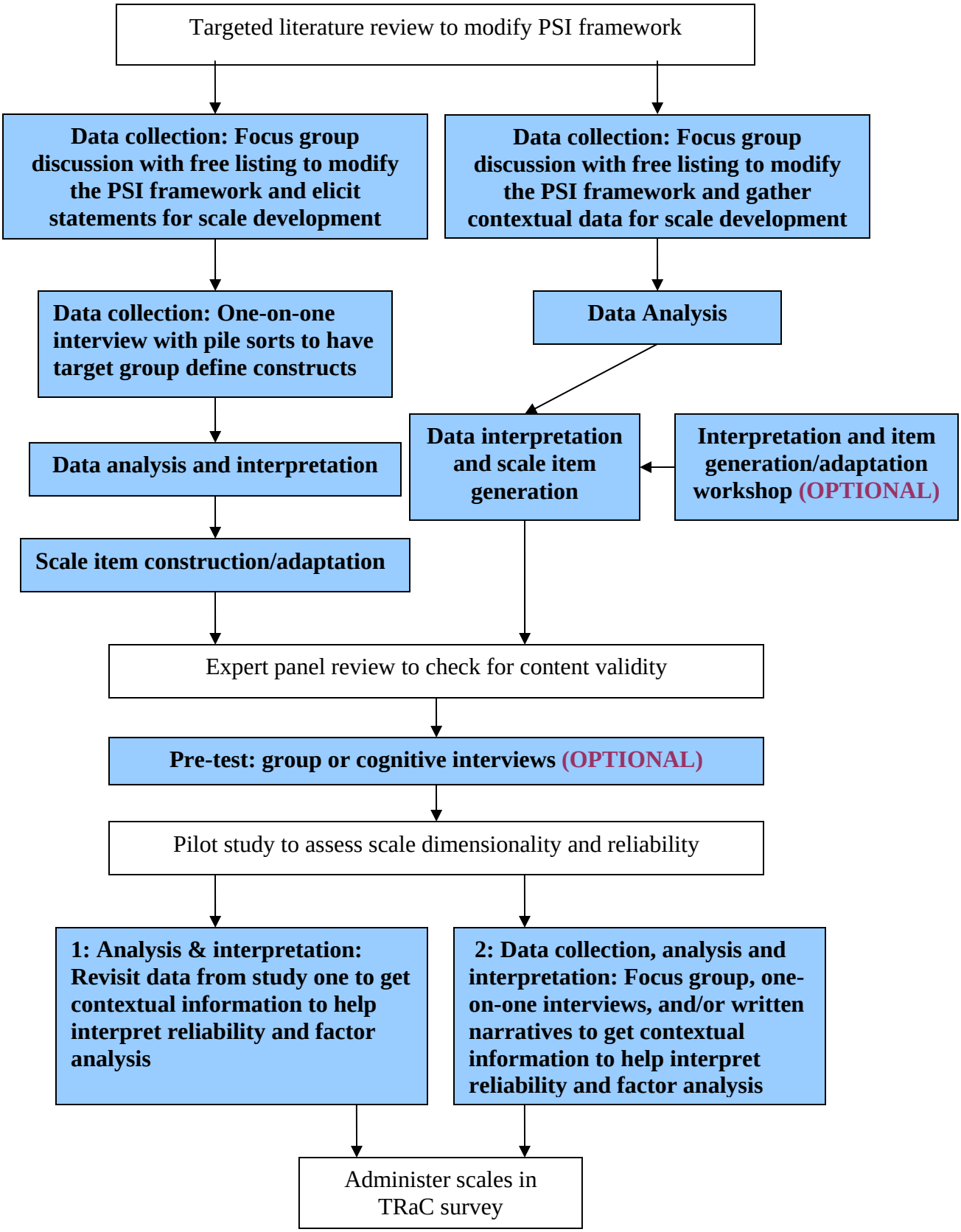


Figure 5: Option III FoQus on Scales Implementation



Advantages and Limitations

The three options for implementation of FoQus on Scales provides for increasingly more input from members of the target population. These phases also gradually increase the amount of resources and technical capacity needed to implement FoQus on Scales.

Option I allows for less input from the target audience than either Option II or III. If workshops are used in Option I the amount of input from the target audience increases substantially. The advantage of starting at with Option I is that it takes less time and resources since it is possible to use only one set of qualitative studies. A moderate level of technical capacity is required from staff involved in data collection and higher level data analysis skills are needed from staff conducting analysis, defining determinants, generating items, and facilitating workshops.

In Option II members of the target audience do not only provide input but are actually involved in defining determinants and generating scale items. Target population members generate and define statements and through the pile sorts combine these statements into determinants. This information is used to modify the PSI Behavior Change Framework. Under Option II more significant changes to the PSI Behavior Change Framework are possible. The advantage of Option II is likelihood the data will increase the validity of determinants and the explanatory power of relationships. The approach is also more applicable for theory-generating research, and may result in a new or adapted version of the current PSI Behavior Change Framework Option II also requires more resources and time as two sets of studies that have to be conducted one after another are needed. Moderators must have advanced facilitation skills and understanding of the research objectives. The use of a specific software program (Anthropak) and skills related to the use of this program are also necessary to the analysis of the pile sort data.

Option III allows the most input from the target population. The chances of increasing the validity of multi-items scales are very high. Moderate to high moderating and high analysis skills are needed. Increased time and resources are necessary since a total of three studies are required. Additional skills are needed to develop interview guides that address interpretation issues in exploratory and reliability analysis.

Resources and Capacity

Implementing FoQus on Scales will require staff that are trained and additional time and funding (see Table 1). . Recruiters, moderators, note-takers, transcribers and translators are needed for all three options of FoQus on Scales. Interviewers are needed for Option II.

Prior to implementing Option I research managers will receive training through regional FoQus workshops and follow up TA. Training of research managers for Options II and III will take placed through in-country TA from regional researchers and/or consultants. Implementation of required activities for Options I and II should take about three months to implement. Implementation of Option III will take approximately five months.

Table 1: FoQus on Scales Implementation Resources

Activity	Staffing requirement	Training	Time	Estimated cost
Phase I:	• Recruiters • Moderators • Note-taker • Transcribers • Translators	FoQus regional training Follow up TA	14 weeks	10,000 to 20,000 [1]
Phase II:	• Recruiters • Moderators • Interviewers • Note-taker • Transcribers • Translators	TA only	14 weeks	NA
Phase III:	• Recruiters • Moderators • Interviewers • Note-taker • Transcribers • Translators	TA only	23 weeks	NA

[1] **Based on cost for 2006 FoQus on Scales Haiti Youth HIV/AIDS study**

Global Implementation

Option I is designed for implementation in country programs with sufficient financial resources and technical capacity by the end of 2008. FoQus on Scales is suggested for country programs that have research funding for three to five years. Training of trainers (TOT) activities and regional training workshops for Option I will take place in second half of 2007 and the first half 2008. TOT for regional researchers will include: 1) readings and small group working sessions prior to the research workshop; 2) facilitating regional workshop sessions; and, 3) modeling exercises after the regional workshop. The modeling exercise will involve working with Project Scale PIs to implement one aspect of FoQus on Scales.

Project Scale PIs will assist regional researchers implement one Option I in FoQus on Scales study. As part of this process, project Scale PIs will provide TA over e-mail and make one TA trip dedicated to data collection, data interpretation or generation or scale revision. Whenever resources allow qualitative methods consultants will be involved in these modeling exercises.

Regional trainings will not be conducted for implementing Options II and III. Training for these options will take place through regional researcher TOT training sessions and modeling exercises with Project Scale PIs and consultants with specialization in these areas.

Appendix
Definitions of Terms

Bubbles refer to a generalized component of the PSI Behavioral Change Theoretical Framework, which can be further defined as either a construct or a determinant.

Determinants refer to complex concepts regarding attitudes, beliefs, attributes which can measured and organized into a theoretical framework. Within a theoretical framework, determinants are related to one another and to one or more behavioral outcomes.

Dimensionality refers to the homogeneity of items within a scale and resultant subscales (unidimensional scales refer a single scale and multidimensional scales refer to a scale with two or more subscales).

Emic refers to the assumption that there is little overlap across contexts and cultures and focuses on the adaptation and/or development of specific concepts and measures through data collected from members of the target audience

Etic refers to the assumption that the same concepts are relevant across different contexts and that they can be defined and measured the same way across sites or cultures.

Free-listing is a method whereby focus group participants are asked to list beliefs, attitudes, perceptions within a certain topic area, e.g., reasons why adolescents don't use condoms. These participants are then asked questions to elicit further explanations and contexts for the listed statements.

Interview guides are a set of questions organized by topic for inclusion in a qualitative interview (focus groups, individual interviews). The guide differs from a survey or structured questionnaire as items on the guide may be asked in different order than they appear, and questions may be added or dropped based on data collected.

Pile sorts are a method whereby participants are asked to sort a set of statements by some common denominator. Pile sorting provides data which can be analyzed through aggregate proximity matrices, which provide a visual representation of 'clusters' of items. Quantitative

analysis of these data can further assist in determining significant differences or similarities between sub-populations, e.g., males and females in their clustering of the items.

Reliability is the extent to which an instrument will produce the same results when used two or more times and reflects the accuracy of the scale.

Scales are a group of three or more items or statements, which are designed to measure a single construct, e.g., condom self-efficacy, which can be scored and measured as a single variable for quantitative analyses. Multi-item scales are likely to be more reliable over time than single items because of the scope and precision of information gathered. The variety of statements used to get at different aspects of a construct increases the **scope.** The fact that scales have different response options and use several items to measure a concept allows for a larger variation in terms of the choices that respondents have and thus more **precision** in responses.

Target audience/Target population is the group or population which will be the focus of the research and the intervention, or marketing program.

Validity is the extent to which an instrument measures what you would like it to measure. Validity refers to the extent to which a scale accurately measures a particular construct.

References

Berry, J. W. & Dansen, P. R. (Eds.). (1974). *Culture and Cognition: Readings in Cross Cultural Psychology.* London: Methuen.

Borgatti, S. (2002). A Statistical Method for Comparing Aggregate Data Across A Priori Groups. *Field Methods 14*(1): 88-107.

Bradley, F., Wiles, R., Kinmonth, A-L, Mant, D, & Gantley, M. (1999). Development and Evaluation of Complex Interventions in Health Services Research: Case Study of the Southhampton Heart Integrated Care Project (SHIP). *British Medical Journal 318*(7185): 711-715.

Camfield, L. & Ruta, D. (2007). 'Translation is Not Enough': Using the Global Person Generated Index (GPGI) to Assess Individual Quality of Life in Bangladesh, Thailand, and Ethiopia. *Quality of Life Research. 16*(6):1039-51.

Chapman, S. (2004). Evaluating Social Marketing Interventions. In M. Thorogood & Y. Coombes (Eds.), *Evaluating Health Promotion: Practices and Methods* (93-110). New York: Oxford University Press.

Chapman, S. & Coombes, Y. (2003). Project TRaC: Tracking Results Continuously. *Concept Paper*. Washington, DC: Population Services International, Research Division.

Chen, H.C., Chen, M.L., Lotus Shyu, Y.I., & Tang, W.R. (2007). Development and Testing of a Scale to Measure Caregiving Load in Caregivers of Cancer Patients in Taiwan, the Care Task Scale-Cancer. *Cancer Nursing 30*(3): 223-231.

Chwalow, A.J. (1995). Cross-Cultural Validation of Existing Quality of Life Scales. *Patient Education and Counseling 26*(1-3): 313-318.

Cole, K., Waldrop, J., D'Auria, J., & Garner, H. (2006). An Integrative Research Review: Effective School-Based Childhood Overweight Interventions. *Journal for Specialists in Pediatric Nursing, 11*(3), 166-177.

DeVellis, R. F. (2003). *Scale Development: Theory and Applications (2nd ed.)*. Thousand Oaks, CA: Sage Publications.

Fishbein, M. (2000). The Role of Theory in HIV Prevention. *AIDS Care 12*(3): 273-278.

George, J., Mackinnon, A., Kong, D. C., & Stewart, K. (2006). Development and Validation of the Beliefs and Behavior Questionnaire (BBQ). *Patient Education and Counseling 64*(1-3): 50-60.

Graves, K.D. (2003). Social Cognitive Theory and Cancer Patients' Quality of Life: a Meta-Analysis of Psychosocial Intervention Components. *Health Psychology 22*(2): 210-219.

Greenhaigh, T., Chowdhury, M., & Wood, G. (2006). Story-Based Scales: Development and Validation of Questionnaires to Measure Subjective Health Status and Cultural Adherence in British Bangldeshis with Diabetes. *Psychology, Health, and Medicine 11*(4): 432-448.

Kim, N., Stanton, B., Li, X., Dickersin, K., & Galbraith, JJ. (1997). Effectiveness of the 40 Adolescent AIDS-Risk Reduction Interventions: a Quantitative Review. *Journal of Adolescent Health 20*(3): 204-215.

Lee, L., Avis, M., & Arthur, A. (2007). The Role of Self-Efficacy in Older People's Decisions to Initiate and Maintain Regular Walking as Exercise – Findings from a Qualitative Study. *Preventive Medicine, 45*(1), 62-65.

Patel, D. & Chapman, S. (2004). PSI Behavior Change Framework "Bubbles": Proposed Revision. *Concept Paper*. Washington, DC: Population Services International.

Patel, D. & Chapman, S. (2005). The Dashboard: A Tool for Social Marketing Decision Making. *Concept Paper*. Washington, DC: Population Services International, Research Division.

Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods. 3rd Edition*. Thousand Oaks, CA: Sage.

Simmons, V. N, & Brandon, T.H. (2007). Secondary Smoking Prevention in a University Setting: a Randomized Comparison of an Experiential, Theory-Based Intervention and a Standard Didactic Intervention for Increasing Cessation Motivation. *Health Psychology 26*(3): 268-277.

Spector, P.E. (1992). *Summated Rating Scale Construction: An Introduction*. Newbury Park, CA: Sage Publications.

Stanton, B., Black, R., Engle, P., & Pelto, G. (1992). Theory-Driven Behavioral Intervention Research for the Control of Diarrheal Diseases. *Social Science and Medicine 35*(11): 1405-1420.

Tsai, J.H. (2002). Psychometric Evaluation of the Demands of Immigration Scale with Taiwanese-Chinese Immigrants: a Pilot Study. *Journal of Advanced Nursing 39*(3): 274-280.

van Empelen, P., Kok, G., van Kesteren, N.M., van den Borne, B., Bos, A. E., & Schaalma, H. P. (2003). Effective Methods to Change Sex-Risk among Drug Users: a Review of Psychosocial Interventions. *Social Science and Medicine 57*(9): 1593-1608.

Willis, G. B. (1999). *Summated Rating Scale Construction: An Introduction.* Newbury Park, CA: Sage Publications.