Sue Ellen Hansen, Hyun Jung Lee, Yu-chieh (Jay) Lin, and Alex McMillan, 2016

Introduction
Guidelines
1. Ensure that technical instrument design is appropriate to the method of administration and the target population.
2. Develop complete technical instrument design specifications for the survey instrument, specifying cultu specific guidelines as necessary.
3. Develop language-specific guidelines for the survey instrument as necessary.
4. Develop interface design rules for computerized survey applications, and for self-administered paper instruments.
5. Establish procedures for quality assurance of the survey instrument that ensures consistency of design, adapting evaluation methods to specific cultures as necessary.
6. Consider all possible formats and layouts, particularly when a survey is self-administered on devices pr to the respondent or administered on the respondent's personal device or devices that respondents can acc complete surveys in a public setting.
7. Maintain complete documentation of source and target language or culture-specific instruments, includ specification and design guidelines, and provide comprehensive summaries of the same for data dissemina analysis.
References
Appendices: A | B | C | D | E | F | G | H | I

Introduction The technical design and implementation of a given survey instrument can be viewed separately from questionnaire design (see Questionnaire Design). Instrument technical design focuses less on questionnaire content more on the design of the actual survey instrument that delivers the questionnaire content. In this sense, technical de includes the format, layout, and other visual aspects of the presentation or context of survey questions. In some insta survey design, questionnaire design, and technical design overlap. Mode decisions, for example, may shape the tech format of questions as well as their wording.

These guidelines will use the more general terms 'survey instrument' or 'instrument' when describing procedures or f that apply to the technical design of both paper and computerized instruments, and the term 'application'—which su the need for at least some programming—when discussing procedures for the development of computerized instrum When there is a need to distinguish between types of computerized instruments, such as computer-assisted (compute but not necessarily accessed via the Internet) and Web instruments, reference will be made to the mode-specific type computerized survey.

Study design decisions related to mode have an impact on instrument technical design requirements (see Study Desi Organizational Structure and Data Collection Implementation: General Considerations). Such decisions include whe survey is to be self-administered or interviewer-administered and whether it is to be administered on paper or compu If the survey is self-administered, a decision must be made about whether it should be a paper or a computerized sur a computerized survey, depending on whether it is a computer-assisted self-interviewing (CASI) instrument or a We instrument, there may be effects on the programming costs and the computer user interface—that is, what responder on the computer screen and how the computer interacts with the respondent.

If the survey is interviewer-administered, decisions may have to be made about whether the instrument should be computerized or on paper and whether it should be in-person or by telephone. There may be technical design consid associated with each of those decisions, as discussed below.

If the survey is to be administered on the respondent's personal device, the design and layout must adapt to the vari possible formats. Web instruments need to be developed with the assumption that respondents may complete them c computers, smartphones, or tablets.

Study design also involves decisions about data output, coding, and data documentation (see Data Processing and St Adjustment and Data Dissemination). Thus, design decisions may have an impact on technical instrument design, pr affecting survey implementation in three ways:

1. How easy it is for an interviewer or respondent to use the survey instrument and to provide appropriate response 'usability' of the instrument). This can help minimize user burden.
2. How easy it is to program a computerized instrument and to test it.
3. How easy it is to code, output, analyze, and document survey data.

An instrument's technical design can impact measurement error, including error resulting from cognitive processing effects, and interviewer effects. In the case of multinational, multicultural, or multiregional surveys, which we refer '3MC' surveys, problems in each of the different technical implementations of survey instruments may lead to differ errors. For instance, local implementations could increase interviewer or respondent burden that will lead to cognitiv processing errors or even terminated interviews. Poor design of survey instruments may also increase nonresponse e the levels of the household or respondent (unit nonresponse) or the survey question (item nonresponse). These guide intended to help researchers at coordinating centers and individual survey organizations of 3MC surveys understand instrument technical design requirements and how to approach creating instrument technical design specifications, v at the centralized level, the local level, or both. Study design may dictate how much is specified at the central level a much is left to local survey organizations. While there may be flexibility in this regard, it is important that technical across local surveys leads to survey data that can be compared across cultures and that does not contribute to measu error. For example, question labels should be consistent across survey implementations. Differences across cultures lead to adaptations in technical design across surveys. In such cases, it is important to document the reasons for adap In general, the layout in the source questionnaire should be preserved in subsequent translated versions. That is, the translated version and the original should look exactly the same except for the words. Some examples from previous in the European Social Survey (ESS) where there were differences in layout between the translated version and the s version include: (a) show cards containing the start of the response sentence when the original did not, or vice versa show cards putting the answer codes in boxes, omitting the numbering of the categories, or drawing arrows to indica end points, where that was not the case in the original; and (c) survey items that were formatted as single questions, with their own answer scale, rather than formatted as batteries of items . All of these types of deviations can contrib measurement error. Further examples of areas to mitigate differences in instrument design between the source and ta questionnaires are detailed in the guidelines below.

↑ Ba

Guidelines Goal: To minimize measurement error, nonresponse error, and respondent and interviewer burden due to technical instrument design, and thus maximize the amount of valid and reliable information obtained within an allo budget and time and at the specified level of precision.

↑ Ba

1. Ensure that technical instrument design is appropriate to the method of administration and the target popu
Rationale The design requirements for self-administered surveys differ from the design requirements for interviewe administered surveys. Self-administered surveys have no interviewer to help repair misunderstandings, and there is opportunity to 'train' respondents on how to respond. Computerized instruments, which involve human-computer int call for design features that facilitate such interaction. The design requirements for computerized instruments also d from the design requirements for paper instruments. Interface design rules (see Guideline 4) and quality assurance procedures (see Guideline 5) for self-administered, interviewer-administered, paper, or computerized surveys should developed in advance, and implemented and documented (see Guideline 6) throughout the data collection process. T characteristics of the target population (education, survey experience, literacy, computer literacy, etc.) should influe instrument design decisions. Self-administered surveys are useful only if administered to populations with high liter rates; computerized surveys require target populations with familiarity with computers, or situations in which data c can be facilitated by trained interviewers. Technical instrument design specifications should include as many culture

(see Guideline 2) and language-specific (see Guideline 3) guidelines as possible. For mode selection and its specific considerations, please see Study Design and Organizational Structure. ***Procedural steps***

1.1 Determine whether to develop an interviewer- or self-administered instrument and whether to use a paper o computerized instrument. Some points to consider from a technical design standpoint are:

1.1.1 Interviewer- vs. self-administered instrument:

- Self-administered instruments including CASI, audio computer-assisted self-interview (A-CASI), videc computer-assisted self-interview (video-A-CASI), mail, or Web may lead to better data quality for surv extremely sensitive questions (drug abuse or sexually deviant behavior, for example) . However, there c cross-cultural differences. See Data Collection: Face-to-Face Surveys and Study Design and Organizati Structure for more in-depth discussion.
- Self-administered instruments should make it easy for respondents to recognize instructions (such as "S one") and to read questions, navigate correctly through the instrument, and enter responses . Instruction appear where they are needed, such as "Start here" before the first question and response entry instructi "Tick all that apply") after the question text, and recording responses should be displayed in the order o likely occurrence. In addition, instructions to skip questions should be avoided or used sparingly in pap administered instruments because they can lead to response errors.
- Self-administered components can be combined with interviewer-assisted components of surveys. An interviewer-administered instrument would be better when there is a need to explain concepts and probe responses or when sections of the interview are sensitive (see Guideline 7).
- Interviewer-administered instruments make it easy to perform required tasks in the order in which they expected to be performed. For example, interviewers' instructions such as referring to show cards or otl reading questions, providing definitions, probing responses, and recording responses should be displaye order of respondents' likely occurrence. This is true in both paper and computer-assisted instruments.
- Interviewer-administered computerized instruments may lead to higher data quality in long and comple surveys (for example, providing consistency checks or preloaded information throughout the whole inst or those with embedded experiments (for example, randomizing the order of questions or response opti
- Whether interviewer- or self-administered, the instrument technical design should help to minimize the placed on interviewers and respondents, which increases as instruments increase in length and complex

1.1.2 Paper or computerized instrument:

- Paper instruments may be less costly to develop, but entail additional data entry costs after data collecti may affect the timeliness of data dissemination (see Data Dissemination).
- Computer-assisted and Web instruments require programming, but Web surveys generally are less costl because of the lack of interviewer costs, and don't necessarily require professional programmers for bas programming. On the other hand, if not programmed well, they may introduce higher costs during data processing.
- Some countries or regions may not have the professional expertise in place to do computerized surveys can be infrastructural constraints in some contexts that make it difficult to collect data with telephone o survey instruments (e.g., the lack of sufficient telephone or Internet penetration, or the lack of an adequ frame) (see Sample Design).
- Paper instruments should be less complex than CASI and Web instruments in order to minimize respon burden, but still allow for embedded experiments.
- Web surveys should be designed to be viewed on phones, tablets, and any other mobile devices in addit the traditional computer components.

1.2 Determine whether there are additional design considerations related to characteristics of members of the ta population, such as children, men, the elderly, or the visually or hearing impaired . Ensure that all such consid are reflected in the technical specifications for the survey instrument (see Guideline 2).

1.2.1 Computerized instruments with images to show response options or color-coded keyboards to enter re options can be alternative design solutions for populations with low rates of literacy (see Appendix F) or computer usage experience (see Appendix G).

1.2.2 Instrument designs for interviewing multiple people within the same household and using the same ins may need a customized interface with specific instructions to accommodate the flow of the interview for computerized and paper instruments (see Appendix I).

Lessons learned

1.1 The use of survey computer-assisted methods can help camouflage complexity and facilitate the tailoring of instruments to special populations. For example, describe the results from a number of Dutch surveys of speci populations using computer-assisted interviewing and self-administered components, in which instrument des administration were tailored to target population needs. For example, a simple but attractive screen layout was survey grade school children. In addition, students only needed to use simple keystrokes to answer questions a could stop temporarily when they felt tired. As a result, item nonresponse was reduced compared to a paper questionnaire. They concluded that well-designed computer-assisted instruments both improve the quality of minimize the burden experienced by respondents and interviewers.

1.2 Use of computerized instruments is possible even with low literacy rates. For example, describe a study of educated men and women in rural South India that experimented with an A-CASI instrument and color-coded response options. Although only 10% of participants had ever used a computer before, 80% stated that the ins was user-friendly and felt comfortable responding to sensitive questions.

1.3 Study design should consider the potential measurement effects that may arise from differences in methods survey administration. A review of paradata from the ESS and the International Social Survey Programme (IS revealed some differences in results across countries between those that implemented paper self-administered by mail and those that used interviewer-assisted self-administered surveys or face-to-face surveys.

↑ Ba

2. Develop complete technical instrument design specifications for the survey instrument, specifying culture-s guidelines as necessary. ***Rationale*** Technical instrument design specifications guide the formatting or programming survey instrument or application. They ensure design consistency across culture-specific instruments (to the extent p and facilitate post-production data processing, harmonization, documentation, and analysis (see Data Processing and Statistical Adjustment and Data Harmonization). The following should be taken into consideration:

- The formatting of information and areas for recording responses.
- The formatting of specific text elements such as question text, response scales, and respondent or interviewer instructions.
- The formatting of specific question and response types.
- The linking of survey instrument information and variables in a dataset and the documentation of the instrumen dataset.
- Rules for the use of numbers, color, graphics, images, maps, and icons.
- Specifications for how question formats may differ across different data collection modes.

A coordinating center's specifications should clearly outline the source questionnaire and its content, provide rules f formatting the survey instrument, and suggest appropriate instrument design adaptation strategies for other cultures. agencies may have to adapt specification rules further to adhere to local standards for design of instruments and staf training and other organizational constraints. Any such adaptations should be documented. Note that similar guideli necessary for a data entry application (see Data Processing and Statistical Adjustment). Generally, this guideline is r to formatting of elements in either paper or computerized instruments, although a few may relate to only one or the

other. Guideline 4 adds guidelines that are relevant specifically to computerized applications and their interface desi to self-administered paper instruments.

Procedural steps

2.1 At the beginning of the instrument specifications, provide an overview of the survey instrument, including t of core chapters and required placement of culture-specific chapters (see an example in Appendix C). Make s the formatting adapts for cultural differences (see Adaptation). For example:

2.1.1 Differences in the formatting of information and areas for the recording of responses , including:

- Dates and times (e.g., 24-hour vs. 12-hour clocks).
- Calendar, holidays, and start-of-week.
- Numeric formatting (e.g., thousands, million, billion, and decimal separators).
- Names and addresses (e.g., last name first or second).
- Telephone numbers (e.g., with or without local prefix).
- Currency and monetary values (e.g., placement of currency symbol and negative sign).
- Sizes and measurement (e.g., metric vs. imperial units, Celsius vs. Fahrenheit, clothing sizes, etc.).

2.2 Provide rules for the consistent formatting of specific text elements, such as question text, response scales, respondent or interviewer instructions, and so on. These might include, for example :

2.2.1 Display question text more prominently than response options.

2.2.2 Distinguish interviewer or respondent instructions; for example, using a smaller font of a different col italicizing in parentheses.

2.2.3 Place text elements where and in the order they are needed based on interviewer or respondent task de for example, in an interviewer-administered instrument, a show card instruction precedes question text, a probe instruction follows it.

2.2.4 Evenly space response options in a scale, grid, or table, so that they appear of equal weight or promine

2.2.5 Underline question text that should be emphasized.

2.3 Provide rules for the formatting of specific question and response types (for example, open- vs. closed-ende other information. Also include examples for each rule; these may include:

2.3.1 Enumerated or fixed choice response options (e.g., 1=Female, 2=Male).

2.3.2 "Tick[/Check/Select] all that apply" (e.g., additional options like "All[/None' of the above" should be a respondents to check/select).

2.3.3 Short or fixed-length text (e.g., the maximum number of words should be listed for respondents to pro answers).

2.3.4 Open-ended text (e.g., the maximum number of words should be provided as needed).

2.3.5 Numeric responses (e.g., for computer-assisted instruments, the range check should be provided and b for quality control).

2.3.6 Response entry masks (e.g., "__/__/____" for dates).

2.3.7 Multi-part questions and question series; for example:

- Day/month/year (e.g., either numeric or text value examples, like "01" or "January," should be provided
- Address/contact information (e.g., instruments should list address info to levels like country, state, coun street, zip code, etc.).
- Demographics question sets.
- Amount-per-unit (e.g., income per day/week/month/year).

2.3.8 Randomly ordered questions, response options, or sections.

2.3.9 Response scales.

- Fully-labeled scale.
- Partially-labeled scale.
- Roster or grid. Rosters are tables used to collect various information in columns about entities in rows; example, gender and age (columns) about persons in a household (rows). Grids are often used for scale (columns) on a number of items (rows).

2.3.10 Text fills (variable question text); for example, question text may vary based on size of household—"y respondent in a single-person household, and "you and your family living here" for a household with mu persons.

2.3.11 Visual or contextual indicators that help respondents or interviewers understand where they are in a qu series (for example, indicating above or beside a series of questions which household member, vehicle, o of income they are discussing).

2.3.12 Progress indicators (i.e., a visual indicator of where the interviewer or respondent is in the instrument survey progresses; applicable only for electronic instruments).

- Progress indicators are speculated to reduce breakoffs, but added graphics associated with the use of a p indicator increases download time .

2.3.13 Question-level help for use as necessary by the interviewer (question-by-question objectives, includin definitions) in paper or computerized surveys.

2.3.14 Validation or consistency checks and post-collection edits. For paper instruments, these should be note instrument technical design specification for use in post processing. In computerized surveys with progra consistency checks that occur during the survey interview, there is a distinction between a *hard* consister check (the interviewer or respondent cannot continue until an inconsistency is resolved) and a *soft* consis check (the interviewer or respondent may continue without resolving the inconsistency).

2.4 Add information to the instrument specifications that facilitates recording responses, the linking of survey instrument information and variables in a dataset (a data dictionary), and the documentation of the instrument dataset (traditionally called a codebook) (see Data Dissemination and Appendix C). For example, specify:

2.4.1 How questions are identified in the dataset (variable names and labels) and how response categories ar numerically represented and labeled (value labels).

2.4.2 Open question formats; consider the amount of space needed to provide for responses, which may diff languages.

2.4.3 Pre-coded response options. If necessary, specify international standards for code numbers and classifi such as occupation, language, country of origin, and religion (for example, specifications for the ESS sta

codes for respondents' language(s) are based on the ISO-639-2 code frame, but use alphanumeric codes i dataset).

2.4.4 Code number conventions (e.g., Yes=1, No=5; Yes=1 or No=2; or No=0, Yes=1). Note that code numb generally not shown in self-administered questionnaires. Yes=1 and No=5 is sometimes used instead of ` and 2=No to minimize error in interviewer-administered surveys. This is because the number 5 is farther from the number 1 than the number 2 is on a computer keyboard; thus, 2 (No) is less likely to be pressed the interviewer means to press 1 (Yes).

2.4.5 Categories for missing data categories, such as,

- Not applicable (does not apply to the respondent; question not asked based on prior answer).
- Refusal (respondent refused to answer question).
- Don't know/can't choose.
- No answer (interviewer or respondent did not provide response, including due to errors in computerized instrument programming). Note that interviewing, coding, or statistical software may constrain labels u create survey datasets. Specifications should indicate the values required in the final datasets and in fina documentation (the codebook).

2.4.6 Data input formats, including scales that use metaphors (such as ladders or thermometers).

2.4.7 Interviewer or respondent instructions.

- Respondent show card instructions.
- Routing (skip or filtering) instructions.
- Response format or data entry instructions.
- Question level flag or mark should be added if the question-level Q-by-Qs information has been prepar the interviewer or respondent to use for understanding questions asked.

2.4.8 Universe statements; that is, metadata that indicates a question or question group was asked of a speci group of the survey population (e.g., "Universe [for this question]: Women aged greater than or equal to years").

2.4.9 Variables to construct or recode during post-production.

2.5 Provide rules for the use of numbers, color, graphics, images, maps, and icons.

2.5.1 Ensure that numbers used in response scales visible to respondents do not have specific implications i cultures. For example, some numbers are considered unlucky in some cultures, such as the number thirte United States.

2.5.2 Ensure that colors used in instruments do not have any negative connotations in specific cultures. Colo different meaning across cultures, and research has found there are cultural differences in color preferen choice of colors should be validated by experts on particular cultures . This may involve harmonization t of "culture-neutral" colors across instruments or adaptation of some colors across instruments as necessa example,

- Red in China means happiness while it means danger in the Western countries, as well as in Japan .
- White, black, all shades of gray, all shades of blue, and a light yellow are preferentially used internatior However, be aware of any association of specific colors with political groups in some countries.

2.5.3 Ensure that any maps used are drawn to scale.

2.5.4 Ensure that images are displayed using comparable typographical units across survey implementations

2.5.5 Ensure that graphics, images, and icons convey comparable meaning across cultures and do not have r connotations in specific cultures, or adapt them as necessary.

2.6 If using multiple data collection methods, include specifications for how question formats would differ acrc methods. For instance, a survey may be interviewer-administered in multiple modes (paper and computerized person and by telephone); it may be self-administered in two modes (Web and mail); or it may be self-admini multiple modes (computer-assisted, paper, and Web). For example:

2.6.1 A computer-assisted self-interviewing (CASI) screen might have only one question and input field per or have questions with same response scales per screen (to minimize respondent burden), whereas an interviewer-administered computer-assisted screen might have multiple questions and multiple input fiel

2.6.2 Self-administered instruments may be developed without response codes (the respondent clicks on a r option, clicks on a radio button, or checks a box), whereas some computer-assisted personal interview (C surveys may require numbered response options for the entry of responses, if numbers are the only possi of input.

2.6.3 Software constraints may also necessitate alternate specifications, for example, if different types softw used for Web and computer-assisted telephone interviewing components.

2.7 Based on the guidelines specified above, as well as the interface design and paper instrument guidelines tha prepare a survey instrument specification with all survey contents for the instrument as well as a data dictiona which represents the contents of the survey dataset. Also specify the codebook metadata before data collectio

Lessons learned

2.1 Seemingly small differences in instrument design across cross-cultural surveys can influence responses acrc cultures. Scales that are not formatted consistently, response options with misaligned check boxes, differences relative amount of space allowed for open responses, and differences in the physical placement of follow-up c have been shown to lead to missing data or unusual response distributions across surveys . For example, in th ISSP there was a question on subjective social stratification. Respondents in nine countries were asked to rate themselves on a scale from 1 to 10 (top to bottom). Respondents in all countries tended to rate themselves in t middle, and a small proportion of respondents rated themselves in the bottom. However, the Netherlands had the middle, compared to 72% to 84% in other countries, and had 37% in the bottom, compared to 6% to 24% countries. Dutch respondents did not have such a distinctive distribution on other social inequality measures. examination, it was found that the Dutch translation was comparable to English, but the visual display of the s differed (see Appendix D).

2.2 On the other hand, cultural customs and norms may require using different graphic images, icons, colors, et example, in 2007, the ISSP allowed countries to use different graphics for an ideal body shape question. See A D for images used in the Austrian and Philippines questionnaires.

2.3 The layout of scales should not deviate from the source questionnaire, e.g. a horizontal scale should never b changed into a vertical scale. Likewise, the order of response categories should not be reversed, e.g. "extreme happy"–"extremely unhappy" should not become "extremely unhappy"–"extremely happy". Such changes car contribute to measurement error .

2.4 When underlining is used to emphasize words or phrases to be stressed by interviewers, the emphasis shoul maintained in the target language questionnaire. This may at times mean that a different word or groups of wc need to be stressed if a close translation has not proved possible .

2.5 Hashtags are commonly used on social media platforms such as Twitter, Instagram, Facebook, etc., and mu data extraction tools or Application Programming Interfaces (APIs) have been developed for researchers to ac these data (see Appendix H). Data entry instructions may need to be provided to respondents or social media include a "#" before entering any response, or to record the full response without any space to reduce data pro efforts.

↑ Ba

3. Develop language-specific guidelines for the survey instrument as necessary. ***Rationale*** Different language fe across cultures are important in designing survey instruments. Survey instrument designers should consider both lan and countries or cultures when developing language specifications, since there is no one-to-one match in languages cultures. Some countries share the same language (e.g., English), but may have different language layout systems, a countries use multiple languages (e.g., Belgium and Switzerland). In addition, some countries have more than one s system of writing (e.g., Japan). Therefore, consider any differences across survey implementations in scripts, charac fonts, text directions, spelling, and text expansions when developing instrument technical design specifications . Th important for computerized instruments, since software may need to be configured and instruments reprogrammed t languages in cultures for which they was not originally developed.

Procedural steps

3.1 Provide instrument formatting specifications that facilitate the translation of languages (see Translation: Ov specifying scripts, character sets, fonts, spacing, and so on, for target languages and the programming of comp assisted instruments; formatting guidelines should address aspects of design such as:

3.1.1 Language- and region-specific character sets.

- The International Organization for Standardization (ISO) 8859 Character Set has language-specific gro for example, ISO 8859-1 for Western Europe and ISO 8859-2 for Central and Eastern Europe.

3.1.2 Differences in languages and scripts; for example:

- Japan has one language, but several scripts, which can be mixed.
- China has one official language, Mandarin (*Putonghua*), as well as seven other major languages and ma dialects. Additionally, Chinese may be displayed in either Traditional or Simplified script.

3.1.3 Differences in fonts that support different character sets; in general:

- Avoid complex or ornate fonts.
- Provide interline space to ensure clear separation between lines and to accommodate underlining.
- Provide space to accommodate changes in line heights.
- Provide flexibility in layout of the instrument to accommodate expansion or contraction of text during translation. For example, use a larger font and/or margins for an English instrument, if translating from into other languages would increase the amount of space required for text in culture-specific instrument

3.1.4 Differences across languages in punctuation (e.g., the different question marks in English and Spanish '¿', respectively).

3.1.5 Language- or culture-specific differences in the way characters are sorted alphabetically, including dia (accent marks above or below letters, e.g., 'É'), ligatures (multiple letters treated as single typographical e.g., 'æ,' 'œ,' and 'ß'), character combinations (e.g., 'ch' follows 'h' in Czech), and uppercase and lowercas For instance, the 'Ä' sorts after 'Z' in Swedish, but after 'A' in German. This is important for computerize software that was designed for one type of culture but used in other cultures or countries that sort lists su response options differently.

3.2 Consider differences in text or figure directionality and provide application design specifications that can be to translated instruments with differing text or figure directionality; the three types of text or figure directiona

3.2.1 Left-to-right (Latin, Cyrillic, Greek, Thai, and Indic languages).

3.2.2 Left-to-right and vertical (Chinese, Japanese, and Korean).

3.2.3 Bi-directional (Arabic and Hebrew characters displayed right to left; Latin characters displayed left to

3.2.4 Text directionality applies to displaying images. For example, in Arabic and Hebrew, where the text is from right to left, images are also read from right to left .

Lessons learned

3.1 In Asian countries, vertical text direction is seldom used for survey questions, but it is sometimes used for r options. In the 2006 East Asia Barometer survey, there were differences across countries in the use of vertical Mainland China and Taiwan used vertical text for response options, but Singapore did not. In the ISSP in 200 and China used vertical text. When vertical text was more than one line, they were displayed from left to right Japan, although they were displayed from right to left in mainland China (see Appendix E). These differences both that design specifications need to reflect an understanding of how different Asian countries display text b vertically and horizontally, and that it would be desirable to pretest separately questions that differ across cou

3.2 cautions against administering visual representations to right-to-left readers (of Arabic or Japanese, for exa that are meant to be processed from left-to-right (of English, for example). In studies comparing the results of pictorial inductive reasoning exercise administered to Arabic-educated Nigerian and Togolese high school stu with that of an Austrian calibration sample, researchers found the Arab-educated students exhibited far more using the left-to-right processing format required by the test than the Austrians because Arabic is read from ri left. In a 3MC project in six countries in the Middle East, researchers discovered during the design phase that Arabic-to-English translations uncovered potential differences between how respondents in the Middle East a respondents in western countries visualize and mentally process rating scales .

↑ Ba

4. Develop interface design rules for computerized survey applications, and for self-administered paper instru
Rationale Interface design has an effect on the respondent-computer or interviewer-computer interaction, influences performance, and may affect data quality. Design should not only minimize respondent and interviewer burden and maximize usability, but should also be consistent across survey implementations. Therefore, it is important to provic guidelines for the design of instructions, questions, error messages, and screen elements for computerized instrumer Appendix A for an example of basic design guidelines for computer-assisted surveys). Note that similar rules are ne for data entry applications (see Data Processing and Statistical Adjustment).

Many of the principles for interface design of computerized instruments are also relevant to paper instruments. They as easily address the usability of paper instruments, whether they are for interviewer-administered or self-administer surveys. In the procedural steps below, no distinction is made between computerized and paper instruments if a step apply to both paper and computerized surveys. Where necessary, distinctions are made between computer-assisted a interface design.

Procedural steps

4.1 Establish the key principles for design, which should lead to effective assessment of the quality of design (s Guideline 5). These include:

4.1.1 Consistency.

4.1.2 Visual discrimination among questions and related elements, so that interviewers and respondents quic learn where different elements are located, and thus where to look for what type of element. For example interviewer and respondent instructions may appear in a smaller text, a different font, and/or a color, to distinguish them from the question text.

4.1.3 Adherence to a culture's normal reading behavior for each language and script, based on issues such as directionality (see Guideline 3).

4.1.4 Display of instructions at points appropriate to associated tasks.

4.1.5 Elimination of unnecessary information or visual display of other features that distract interviewers an respondents.

4.2 Provide rules for the layout and formatting of question elements, including:

4.2.1 Question text, which should be the primary focus of a question, and its related information.

4.2.2 Response options, which should have instructions or visual characteristics that convey whether a singl mutually-exclusive response or multiple responses are possible. For example, in computerized instrumer buttons convey that there should be one response, and check boxes convey that there may be multiple re which should be reinforced by an instruction (e.g., "Select all that apply").

4.2.3 Response input fields should convey the length of the response expected. For example:

- A response area is as wide as and has as many lines as the expected length of response.
- The width of an integer response area should be as many number of character lengths wide as the expec input, that is, one character length for a one-digit integer, a two-character length for a two-digit integer,

4.2.4 Instructions, which should appear as expected in relation to task demands; for example, a reference to respondent booklet or show card should appear before question text, and a probe or data entry instructio question text.

- Layout can also play a role when deciding on translations for interviewer or respondent instructions. If instruction reads “Please tick one box” (as in the self-completion supplementary questionnaires), the tra for “box” should match the symbol that is eventually used, such as '□' or '○'. Equally, the translation for should match the actual action (tick? mark? touch?), which can depend on whether the questionnaire is computer- or paper-based .

4.2.5 In computerized instruments, the interface should facilitate accessing online help, through clear forma help text and design of navigational aids that facilitate opening and closing help text windows.

4.2.6 Error messages, warnings, and consistency checks in computerized instruments should clearly identify nature of the problem, reflect actual question wording if necessary (e.g., for interviewer probes for more responses), and convey how to resolve the problem (see for examples and for more detailed guidelines o of error messages).

4.2.7 Context markers (for example, instrument section labels, household member numbers, and so on).

4.2.8 Additional information may be required for Web self-administered surveys, such as contact informatic graphic and/or text identification of the sponsoring organization.

4.2.9 In Web surveys, provide guidance on whether to use a paged vs. a scrolling design . Provide rules for l cultural differences; for example, differences in paper sizes for paper surveys. In such cases, provide gui

pagination in order to avoid inadvertent context effects (or example, two related questions appearing tog one page in one country's survey and on separate pages in another).

4.2.10 Provide examples of key question types and elements for all target languages and cultures, and for dif types of administration if relevant (see Appendix A for examples of computerized questions and Appenc examples of paper questions).

4.2.11 Provide examples of correct formatting of elements, for all question types (see Guideline 1) and all lar and cultures (see Appendix A).

Lessons learned

4.1 There is increasing evidence that the visual design of computer-assisted and Web instruments can impact da quality . For example, providing an input box or field that allows entry of 10 numbers with no guidance or ins on input format can lead to poorer data quality than if the survey question more precisely calls for an integer c three digits; for example, instead of "20," "90," or "100" in an entry field with a width of three (___), a Web s respondent enters "40 to 50" in a field with a width of 10 (-_____) can lead to poorer data quality due to entry errors.

4.2 Not providing rules for formatting questionnaires printed on different sized paper can lead to poorer compa of data across countries. For example, in the ISSP one country lost the last item in a scale when copying the s from A4 size paper (8.27" by 11.69") to letter size paper (8.5" by 11") .

↑ Ba

5. Establish procedures for quality assurance of the survey instrument that ensures consistency of design, ada evaluation methods to specific cultures as necessary.

Rationale

As discussed in Guideline 4 above, research shows that instrument technical design can affect data quality in compu assisted or Web surveys positively or negatively. This is also true of paper instruments. Thus, it is important that pre (see Pretesting) of comparative survey instruments includes procedures for assessing the quality of the design of the instrument and adaptations for specific cultures, languages, and modes, and not just the quality of the content. This i the evaluation of the use of color, graphics, images, maps, and icons. As indicated earlier, such evaluation procedure require adaptation across cultures. ***Procedural steps***

5.1 Identify a team with members that have expertise in the evaluation of technical instrument design. Such exp include substantive experts, survey methodologists, linguists, and usability professionals, and should include s with an understanding of response styles across cultures.

5.2 Provide a clear set of instrument specifications and/or a data dictionary for the instrument and culture-speci adaptations (per rules outlined in Guideline 2), which will facilitate testing and assessment of the instruments. documentation would include: question (variable) names and labels; question text; response option values and numeric response formats and ranges, and specifications for the lengths allowed for open-ended question text interviewer or respondent instructions; missing data values; skip instructions; etc. It should enable comparison computerized or formatted paper instruments to instrument design specifications.

5.3 Identify appropriate instrument evaluation procedures for the comparative surveys under evaluation. These more or less extensive based on whether survey organizations in the targeted cultures have previously used sp guidelines, instruments, and survey software. Most questionnaire pretesting tools (see Pretesting) may be used evaluate instrument design as well as questionnaire content and data collection procedures. These include:

5.3.1 Expert review or heuristic evaluation, in which one or more experts evaluates the instrument design ag set of evaluation criteria or heuristics, for example:

- Consistency and adherence to design guidelines.
- Error prevention.
- Usefulness of documentation, definitions, help, error messages, and other feedback to users.
- Ease of navigation.
- Ease of recognition of specific question or instrument elements and actions required.

5.3.2 Review of an instrument, data dictionary, or codebook to ensure adherence to instrument specification naming and labeling of variables and response options. This should include comparison across instrumer data dictionaries for all survey implementations.

5.3.3 Laboratory or on-site tests of instrument design with users or participants with similar characteristics t interviewers or respondents. These are called usability tests when evaluating computer-based instrument they also may be used to evaluate paper instruments. Since culture-specific response styles affect how participants respond to questions about usability every effort should be made to match tester and particip characteristics, language, and cultural background.

5.3.4 If feasible, incorporate methodological experiments on formatting, to assess whether aspects of forma affect respondents differentially across cultures.

5.4 Test instruments locally on top of central testing.

5.4.1 Field instruments that require Internet connection should be tested for connectivity in field situations. offline alternative should be established if there are connectivity issues.

5.5 Collect measures from all instrument evaluation procedures that will lead to informed decisions about quest screen-specific or global design changes that need to be made (see Pretesting). Examples include:

5.5.1 Questionnaire length and section and item timings.

5.5.2 Audit trails for computer-assisted or Web applications, which can include item timestamps, keystrokes actions, and functions invoked. Gathering some of these requires programming that captures information from the respondent's computer (provides sample programming code for capturing such paradata for We surveys).

5.5.3 Behavior codes or event codes based on video or audio recordings that reflect problems using the surv instrument. Such methods are appropriate for both paper and computer-assisted instruments.

5.5.4 Qualitative analyses of cognitive and usability testing.

5.5.5 Heuristic evaluation or expert review.

Lessons learned

5.1 Research has shown that techniques for evaluating the effectiveness of paper materials and computer softwa very well in the evaluation of the design of survey instruments. For example, usability evaluation methods (cc used in the development of software to assess the quality of user interfaces) and traditional pretesting methods conventional pretests, cognitive interviews, and behavior coding can be used to identify instrument design prc as well as problems related to question content.

5.2 Interviewer and participant interaction may need to be considered for usability tests of instruments used in 3 surveys. There is evidence that when an interviewer is from the same culture as participants, interviewers give help, tell more about introductions, and encourage participants more frequently, and participants report more u problems and give more suggestions than when an interviewer is from a different culture . On the other hand, research indicates that when interviewers are from cultures speaking different languages, participants explain about their choices of design elements .

5.3 Incorporating methodological experiments into cross-cultural surveys, whether for experiments on instrume design or other methodological issues, can be difficult to negotiate. It involves agreement of funding agencies central coordinating center (if there is one), and the survey organizations involved. It also requires that clear experimental design specifications are included as part of the development of design specifications prepared f survey organization (see Guideline 2).

↑ Ba

6. Consider all possible formats and layouts, particularly when a survey is self-administered on devices provi the respondent or administered on the respondent's personal device or devices that respondents can access to complete surveys in a public setting (See Study Design and Organizational Structure). ***Rationale*** A self-admin component may be better when a portion of the interview is sensitive, although this varies by social context (see Da Collection: Face-to-Face Surveys for further discussion). When using CASI and A-CASI modes, attention to the det discussed below that facilitate the respondent experience can lead to increased data quality. ***Procedural steps***

6.1 Ensure that there is a good fit between the project and the technological device.

6.1.1 Handheld devices such as personal digital assistants (PDAs) or smartphones may be more appropriate smaller or simpler questionnaires.

6.1.2 An important limitation of PDAs and smartphones is that they are not as suitable for collecting open-e responses .

6.1.3 Particularly with the use of a PDA or smartphone, researchers need to be aware of the size of the devi relative to the interviewer's hand.

6.1.4 Interviewers might lose track of where they are in the sequence of questions and might find it difficult a comprehensive picture of the instrument since they see only one screen at a time.

6.1.5 Moreover, interviewers might find it more difficult to handle qualitative open-ended questions that rec lot of typing verbatim answers. Handheld devices (e.g., smartphones) are not as suitable for collecting op ended responses as are laptops .

6.2 Implement a system of work ownership. All personnel can be assigned a code for database entry, supervisic analysis. Logs can be generated to monitor and control data management and information flow.

6.2.1 Additional attention should be given to non-Latin languages (i.e., Chinese, Arabic, Russian, etc.) whe selecting technology and programming software. Not all software packages can support non-Latin script

6.2.2 Allocate sufficient time to designing and pretesting the electronic questionnaire and to overall testing a debugging the software, or difficulties can arise due to lack of adequate preparatory time . This is particu questionnaires in multiple languages, especially if the survey uses a non-Latin script and/or if the questic lengthy and complex, as it is crucial to ensure that the question flow and skip patterns function correctly using them in the field.

6.3 Consider using paper documents for certain aspects of the survey. For example, interviewers in China using handheld computers reported that it was overly time-consuming to read the full consent form on a small scree

6.3.1 In a public health survey in China, interviewers reported that entering Chinese characters using the handwriting recognizer was too time-consuming and entering Chinese characters with the stylus into the handheld computer was also difficult .

6.4 When using CASI and A-CASI modes, attend to details that facilitate the respondent experience, leading to improved data quality.

6.4.1 Consider disabling the screen saver and power-saving settings on the device so that screens do not go a participant takes additional time to answer a question .

6.4.2 Graphical and/or audio representations of the response process can help guide the respondent through interview. In a survey using A-CASI in India, the entry of a response was marked by the change in the c the corresponding response bar on the screen to grey, along with a 'beep' sound. A “Thank you” screen i the end of the survey .

6.4.3 If a participant did not answer a question after approximately 60 seconds, consider repeating the quest and/or programming additional text can be programmed to appear encouraging participants to answer th in a truthful manner .

6.4.4 If a keyboard is used, it should be user-friendly.

- Keyboard options can be limited to responses (e.g., "Yes," "No," and numbers) and larger color-coded k keys could be used (see Appendix G).
- Additional keyboard shortcuts to replay questions can also be marked.

6.4.5 Text on the computer screen should be large enough to be easily legible for respondents.

6.4.6 In an A-CASI survey in India, neither the question nor the response texts were displayed on the screen ensure privacy and confidentiality for the respondents .

6.4.7 Touchscreens on A-CASI instruments can be particularly helpful for less-educated populations .

6.5 Consider the different types of mobile devices that a respondent may use to complete a survey. For example surveys may be accessed through computers, smartphones, or mobile tablets, and completed on one or more d 'Bring your own device' (BYOD) has become a trend for telephone surveys and surveys can be administered a most convenient for the respondent.

6.5.1 To achieve its cost and quality targets and meet its strategic goals for Census 2020, the U.S. Census B continues to explore the public’s willingness to be enumerated given a BYOD concept in which intervie using their personally owned devices .

6.6 Consider collecting data using Short Message Service (SMS) text, with reminders sent to mobile phones or use of 'apps' for surveys that specifically target respondents that possess smartphones.

Lessons learned

6.1 Do not underestimate the additional time needed for preparation when using technology. In a survey in Burk Faso, researchers reported underestimating the amount of work required to program questionnaires, and as a r failed to maximize the use of some of the available options for input checking and other real-time quality con procedures. Village names, for example, were implemented as a text-entry field, but would have been better a

down list to avoid ambiguities of spelling, etc. Combinations of input checks, plus quality control measures at stage where data were downloaded to portable computers in the field, should have picked up concerns at an ea remediable stage .

6.2 In a Bolivian survey, interviewers reported that longer survey questions disrupted the flow of the interview of extra scrolling time .

6.3 In a Kenya study, A-CASI had much lower rates of missing data than the paper self-administered questionn similar rates to the standard interviewer-administered paper questionnaire. Use of computers in rural populati sometimes met with suspicion and opposition.

6.4 In a Malawi study , reporting for "ever had sex" and "sex with a boyfriend" is higher in the face-to-face (FT than self-administered A-CASI. Instead, reporting about other partners as well as multiple lifetime partners, h is consistently higher with A-CASI than FTF. Overall, the FTF mode produced more consistent reporting of s activity between the main interview and a subsequent interview. The association between infection status and reporting of sexual behavior is stronger in the FTF mode, although in both modes a number of young women denied ever having sex test positive for STIs/HIV in associated biomarker collection.

6.5 Comparisons with alternate administration modes suggest that the audio self-administered questionnaire mo strongly increased reporting of socially undesirable behaviors. Further analyses suggest that when self-admini is combined with the use of earphones, the threat of bystander disapproval (as opposed to interviewer disappr reduced by effectively isolating respondents from their social environment.

6.6 In Kenya, each text message reminder included a quote that was up to 40 characters long and was unrelated topic of the survey (malaria case management), but was designed to be motivating, entertaining, or merely att getting to increase the probability that health workers would read the messages and respond to the survey .

See also Data Collection: Face-to-Face Surveys for further literature on the use of CASI and A-CASI.

↑ Ba

7. Maintain complete documentation of source and target language or culture-specific instruments, including specification and design guidelines, and provide comprehensive summaries of the same for data disseminatio analysis. ***Rationale*** Comprehensive documentation of survey instruments or applications is an essential component documentation and comes into play at all stages of the survey lifecycle (questionnaire development, pretesting, data collection, post processing, and data dissemination and analysis). Complete and consistent rules for specifying and d instruments are important (although not sufficient) to ensuring survey data meet the quality requirements of users (s Survey Quality). Documentation of instrument design specifications also plays a significant role in this regard. In 3M surveys, it also facilitates the assessment of comparability of survey data across cultures. The rapid increase in comp assisted data collection methods makes it increasingly possible to provide well-documented survey data. Based on s design, the study coordinating center, the survey agency, or both would be responsible for maintaining documentati related to technical instrument design. ***Procedural steps***

7.1 Maintain documentation of the rules specified for technical instrument design.

7.2 Maintain documentation of quality assessments of the survey instruments and the outcomes of decisions ma revise the instrument design.

7.3 Maintain specifications for the final source instruments, based on Guidelines 1, 2, 3, and 4 above. These sh include the instrument specifications and data dictionaries developed by the coordinating center and/or survey organizations.

7.4 Maintain alternative specifications for target languages or cultures as necessary. For example, if the source instrument is computer-assisted, but it is necessary to develop a paper instrument for one or more locations, s specifications should be developed for paper instruments.

7.5 Maintain paper and/or electronic copies of all culture-specific instruments or adaptations of instruments to f comparison of technical design across culture-specific surveys.

7.6 Maintain question-level metadata (question text, response options, instructions, text fills, population univers definitions, etc.) in an electronic format to facilitate linking and comparing metadata for all survey instrument eXtensible Markup Language (XML) data files). If feasible, this should be part of a centralized documentatio that links question metadata and formatting with data codebooks for data disseminated. Some computer-assist collection software now makes this possible.

7.7 Provide comprehensive documentation of survey instruments, based on all of the sources of documentation above.

Lessons learned

7.1 Survey instrument design and documentation of design rules and specifications can affect the quality of dat produced and disseminated, and the ability of users to effectively analyze survey data. conducted studies of us "interacting" with statistical data in order to understand how to better meet their needs. In one study, she foun completeness and quality of available question-level survey instrument documentation and metadata affected selection of variables for analysis. In particular, she found that users used a number of mechanisms for identif appropriate variables for analysis, including what they knew about variable naming conventions, how particul questions relate to other questions, and even coding categories, if the question text did not provide enough inf for selection. These findings reinforce the need for clear documentation of technical design guidelines and ins specifications, and for these to be readily available to data users.

↑ Ba

References

{2265844:IY6SFUI8};{2265844:BFMBBCVM};{2265844:KACAY26L},{2265844:9XBCB4S7};{2265844:6862I
{2265844:6862DIBU};{2265844:ER5UKALV};{2265844:7PBBD5H4};{2265844:7PP5G37Q};{2265844:QHSVI
{2265844:7PP5G37Q};{2265844:7PBBD5H4},{2265844:GTGDNFE6},{2265844:IA3KM7AR};{2265844:IA3KI
{2265844:IA3KM7AR};{2265844:HXBDJ5S2};{2265844:IY6SFUI8};{2265844:IY6SFUI8};{2265844:7PBBD5
{2265844:VGYFNW3T},{2265844:7PBBD5H4},{2265844:J3B4G3FY},{2265844:IA3KM7AR};{2265844:7PBE
{2265844:NBGFDWNT};{2265844:GT3JYQM4};{2265844:IY6SFUI8};{2265844:N3E9VA5L};{2265844:5NL2
{2265844:6862DIBU},{2265844:BM4MFNUH},{2265844:QHSVB7BL},{2265844:7PP5G37Q},
{2265844:TGSMWYEV};{2265844:W4EUJIW7};{2265844:CRI4PX4I};{2265844:S2GY4PIR};{2265844:LTRI9
{2265844:MPZDMZKD};{2265844:MBTJKG8M};{2265844:88GV5NSP};{2265844:UA5QYJ3M};
{2265844:7PP5G37Q},{2265844:KF3ARJMY},{2265844:P2T3TNJE};{2265844:UA5QYJ3M};{2265844:HJHK
{2265844:YT6MIWP9};{2265844:YT6MIWP9};{2265844:QN2ZACW3};{2265844:ER5UKALV};
{2265844:QN2ZACW3};{2265844:ER5UKALV};{2265844:TU7J8GCL};{2265844:KTRZZEFY},
{2265844:38A4VDTT};{2265844:ALKR2XW7},{2265844:WPJHHNRC},{2265844:ASKB4GKV};
{2265844:6MMUBUF6};{2265844:4WXIRE35};{2265844:UA5QYJ3M};{2265844:9QT6CDBV};
{2265844:ALKR2XW7};{2265844:RZX3QIHS} apa creator asc 0 40

↑ Ba