

Checklist for Designing Evaluation Surveys

The purpose of this checklist is to guide evaluators in designing surveys or questionnaires that help answer important evaluation questions. If the wrong questions are asked, posed to respondents inappropriately, or the wrong respondents queried, the needed information will not be gathered. The goal is to develop appropriate sampling frames and construct survey questions that can be interpreted in a reliable and valid manner. The guiding principles are alignment of evaluation information needs, constructing appropriate survey items, validation of the questionnaire, construction of valid sampling frames, and planned statistical analyses to summarize the results.

The proposed draft guidelines for this checklist were developed using The Checklists Development Checklist (Stufflebeam, 2002), and the principles described by Rogers (2002). Initial reviews of the checklist, pilot testing, and field testing activities have not yet been conducted. Consequently, the checklist is considered to be in the embryonic stages of development. It is hoped that even in its elementary form, the checklist will provide initial guidance to evaluators in developing items to answer crucial evaluation questions. Users can assist in the further development and validation of this checklist by sending their comments and suggestions to the authors: Larry McKenzie (LardebMck@aol.com), Sarah Heinemeier (sarahhei@mindspring.com), or Amy Germuth (agermuth@mindspring.com)

Pre-Checklist

1. Identify what information is needed.

- Delineate potential sources of information (clients, stakeholders).
- Communicate with stakeholders about what information they would like collected.
- Determine how clients and stakeholders will use the information.
- Know what decisions will be made by the stakeholders using this information.
- Keep a written record of understandings about what information is needed and who will use it.
- Determine the best instrument of methods of data collection. If a survey is the best instrument, continue with this checklist.

Choosing a Sampling Frame

2. Estimate the cost of collecting data.

- Identify the amount of money available to collect and analyze data.
- Consider, discuss, and document issues associated with non-response (since adjusting for non-response increases data collection and analysis costs).

3. Identify the populations of interest.

- If multiple domains within a construct, specify the domain of interest for the construct being studied. For example, if teacher quality is the construct, specify which aspect is of interest: teacher education, teacher interactions with students or parents, teacher assessment, etc.
- Establish the types of decisions that are to be made from the results and their interpretation.

- ❑ Identify the target population, e.g. the complete group of people for whom the survey results are intended to represent.
- ❑ Identify the accessible population, e.g. the group from which the sample will be selected.

4. Choosing the sample.

- ❑ Agree upon the type of inference (statistical or logical) that will be used to make generalizations about the population from the sample.
- ❑ Choose the types of statistical tests (if applicable) to be conducted.
- ❑ Document the types of statistical tests to be conducted and the sample choice and data collection requirements associated with them.
- ❑ Construct the process by which the sample will be selected from the population (e.g., the sample) will be chosen.
- ❑ Choose the sample for further study.
- ❑ Ensure that the sample will be kept either anonymous or confidential.

Aligning Information with Needs

5. Develop the program evaluation questions corresponding to the identified needs.

- ❑ Identify specific information needs.
- ❑ Establish the purpose of the evaluation (i.e., formative or summative).
- ❑ Write evaluation questions to address the information needs.
- ❑ Specify evaluation questions so that responses will provide useable information.

6. Review the program evaluation questions carefully against the decision(s) that need to be made and how the data will be collected.

- ❑ Have relevant stakeholders review the evaluation questions.
- ❑ Have relevant stakeholders confirm that answers to each evaluation question will provide useful data to them as stakeholders.
- ❑ Have relevant stakeholders confirm that intended respondents are able to read, understand, and respond to questionnaires to be constructed.
- ❑ Have relevant stakeholders confirm that intended respondents would be willing to respond to the questionnaires to be constructed.

Developing Open-ended vs Close-end Questionnaires

7. Determine the type(s) of survey questions to be asked.

While this section is divided into open- and close-ended questions, it is important to recognize that often, a combination of both types of questions will provide the most comprehensive data.

Use **open-ended** questions if:

- ❑ No prior knowledge of how the respondents will answer the question exists.
- ❑ The closed-item format of a question may prejudice the response.
- ❑ Free or unaided recall or top-of-the-head responses is wanted.

- ❑ Information about unexpected or unintended side effects of a new program are wanted.
- ❑ Gathering of explanatory information or further elaboration is desired.
- ❑ Numerical findings in a program evaluation or survey need explanation.
- ❑ Potential respondents have the time and inclination to respond.
- ❑ Evaluators don't have time to develop response alternatives found in closed-end questions.
- ❑ Evaluators have the time and budget to gather this type of information.
- ❑ Higher levels of non-response are acceptable.

Use **close-end** questions if:

- ❑ Evaluators and/or stakeholders have sufficient knowledge to develop a set of exhaustive options that can cover the response range of the respondents.
- ❑ The number of questions required to provide the needed information is large.
- ❑ The time available to collect data to provide the needed information is limited.
- ❑ Evaluators have time to develop response alternatives for each question.
- ❑ Potential respondents are not inclined to respond to open-end questions.
- ❑ Higher levels of participant response are desired.
- ❑ Greater reliability is desired when scoring the responses.

8. Write the survey questions to obtain the needed information from the intended respondents.

- ❑ Before developing new questions, conduct an exhaustive search of the literature for pre-existing instruments that are reliable, valid and have been used by other evaluators.
- ❑ Determine and agree upon the type(s) of survey questions to be asked.
- ❑ Know the reading level of respondents and use in constructing questions.
- ❑ Avoid words open to misinterpretation such as likely, possible and probable.
- ❑ Avoid slang words or expressions.
- ❑ Assure that questions are complete or self-contained. In other words, assure that questions do not lead to or refer to other questions.
- ❑ Keep questions as brief as possible.
- ❑ Ensure that each question yields responses that can be used to answer, at least in part, the evaluation question.
- ❑ Ensure that the collection of relevant questions adequately represents the aspect being evaluation.
- ❑ Make sure questions are clearly worded, e.g., words like very, quite, much, most, few, often, several, never, always, and occasionally have been avoided. It is important to note, however, that these words are useful in responses to questions, such as Likert based items.
- ❑ Avoid using technical jargon.
- ❑ Avoid using acronyms, known only to the evaluator.
- ❑ Ensure that questions are specific, i.e., no double-barreled questions--only one variable is addressed in each question.
- ❑ Make sure Likert scale items are not loaded positively or negatively.
- ❑ To avoid a neutral response, assure that forced-choice Likert items have only 4 alternatives.

- ❑ If respondents can truly feel neutral, allow Likert scale items to contain a neutral or undecided category.
- ❑ Use open and close-end questions, where appropriate, to compliment each other.
- ❑ Assure that survey is feasible for implementation; open-ended survey questions are easier to administer via the phone, whereas both open-and close-ended questions are viable for in-person or web participants.
- ❑ Avoid double negatives or other confusing language.

9. Question Development

- ❑ Determine the affective variable(s) to be measured (attitude, interest, values, and self-esteem are four popular ones).
- ❑ Identify the range of affective behaviors and environments to be measured.(Content analysis of open-ended questions and literature reviews are helpful here.)
- ❑ Establish bipolar adjectives for each variable being measured (For example, if one asks whether someone is very satisfied with a product, one should also ask if the product is unsatisfying.)
- ❑ Use more than one bipolar set to measure a single variable (6-8 items are suggested for large domains).

For items to allow “clear and personalized expression of feelings” (Rogers, 2002):

- Avoid factual statements.
 - Avoid statements that refer to the past instead of the future.
 - Avoid statements that might have multiple interpretations.
 - Avoid statements with little or no relevancy to affective variables.
 - Avoid statements that can be uniformly endorsed or non-endorsed (i.e., have low discrepancy).
 - Avoid words such as: only, merely, just, etc.
 - Ensure that statements are relevant to the domain being measured.
 - Use simple, clear, and direct language in short sentences.
 - Include only one domain or variable in any one sentence.
 - Divide statements into positively and negatively worded statements.
 - Ascertain cultural and temporal relevance.
 - Randomize statement order.
- ❑ Read statements for social desirability, acquiescence, and lack of relationship of attitude and actual behavior. These traits are not desirable; revise sentences to remove these problems.

10. Determine the structure of the questionnaire.

- ❑ Group questions designed to obtain the same information or that are referenced to the same object together since it allows for smooth transitions between sections of the survey.
- ❑ Place bio-demographic questions at the end of the questionnaire.
- ❑ Make sure the questionnaire is attractive (e.g., not too much text on one page, provides adequate room for responses if a paper survey, etc.)
- ❑ Make sure the questionnaire is of sufficient font for an easy read (e.g., no less than 10 point font).

Validating the Questionnaire

11. Pilot- and Field-Test the Questionnaire

- ❑ Have knowledgeable stakeholders pilot the questionnaire by answering all of the questions.
- ❑ Have knowledgeable stakeholders comment on each question with respect to clarity and understandability.
- ❑ Have knowledgeable stakeholders comment on whether they feel the intended respondents will be able to read, understand, and complete the questionnaire.
- ❑ Have knowledgeable stakeholders comment on the relevancy and representativeness of the questions for the topic or program evaluation question to which the questions are referenced.
- ❑ Use feedback to modify and improve the questionnaire.
- ❑ Field test the questionnaire with a sample of respondents from the population of interest.
- ❑ Validate the time necessary to complete the questionnaire and include in the directions or cover letter to respondents.

12. Validate the statistical analyses that will be used to summarize the responses.

- ❑ Use item analyses such as Cronbach's alpha or factor analysis to check if questions in a set are internally consistent.
- ❑ Collapse sets of questions whose internal consistency is low into fewer questions so that the analyses takes place at the question level.
- ❑ Calculate and report standard errors of measurement.