

Developing New Data Collection Tools

STEP 4

Step 3 discussed methodology and means of collecting data relevant to indicators of change in order to answer questions about program implementation and effects.

Step 4 provides ideas for developing new data collection tools, primarily questionnaires, to provide specific information for the evaluation.

Review the link between data collection and objectives

Determine relevant data collection tools

Look for model questionnaires

Modify existing questions to suit respondents

Write new questions (See guidelines)

Pre-test questionnaires

4.1 Review the link between data collection and objectives

First, take time to review the purpose of the evaluation.

- What precise information is necessary to evaluate program results:
 - about participants
 - families
 - volunteers
 - community features?
- Is new data needed to fill a gap in existing sources?
- Do you want to obtain similar data from more than one source?

Need an example?

Data from more than one source provide more information about participants and possible program effects. Parents may observe that children are showing stronger social skills and less conflict behaviour at home after taking part in a program. Your confidence in this result would increase if teachers, who did not necessarily know about the program, reported similar improvements in behaviour at school.

- Decide how you can best obtain the information with your resources and the time available, keeping in mind timing, confidentiality, likelihood of cooperation from sources
- A chart to show the need for information, sources and methods will help point to gaps in the process (See R5 & R6 in Appendix.)

4.2 Determine relevant data collection tools

Distinguish between the need for standardized instruments and non-standardized data collection tools such as questionnaires and scales.

Standardized instruments have been developed by specialists to measure psychosocial concepts or qualities, e.g., self-esteem, self-control, sense of belonging, sense of community, family functioning. They have been tested in various versions, some abbreviated, in different languages, settings and populations to determine their level of reliability and validity. Using a standardized instrument conveys some advantages: you can be confident of its findings and may compare them to findings in other comparable populations.

Non-standardized data collection tools, especially questionnaires, are created to measure specific effects of a program, e.g., homework completion, use of spare time, reading habits, recreational skills. They may be borrowed or adapted from other programs, with due credit given, but have no established level of reliability or validity.

No questionnaire is perfect, i.e., every question understood and answered as intended by every informant to provide a completely accurate picture of reality in a given situation. There will always be a certain degree of error in results.

Non-standardized questionnaires have an unknown degree of error and thus results based on them from different times or venues should be compared with caution. They can only suggest similarities or differences among programs or groups and should be supported by other evidence.

Using standardized instruments

If you are interested in investigating effects on the psychological well-being of participants, it may be possible to use one of many standardized questionnaires or instruments used for academic research. Using questionnaires that have been developed, tested and critiqued by others can save considerable time. Online data bases are good sources of information about tests and instruments, many available through university or public libraries. Several published volumes also contain catalogues of tests and questionnaires. Most instruments are copyrighted. It is necessary to contact authors and/or publishers for permissions to use all or part of a standardized instrument. Some may be ordered as kits for a price, others are free with proper attribution given to the author in a credit line on the form. Consider consulting someone experienced in using such instruments, e.g., in research settings such as colleges and universities or larger social service agencies.

Suitability for community programming

- designed to suit specific populations and may not be transferable to other groups, e.g., adult to children, urban to rural, U.S. demographics to Canadian.
- sometimes, questions or specific wording may not be entirely suitable for the group or purpose
- must be used precisely as written and formatted,
 - **using certain questions and omitting others invalidates the reliability and validity of the instrument, making it unsuitable for comparison purposes.**
- Some standardized tests, however, have been used for decades and proven in a variety of population groups.
- It is possible to add extra questions to a standardized instrument without compromising the instrument's use for comparative purposes, if you analyze the instrument questions separately from the additional questions.
- It may be helpful to borrow ideas and phrasing from questionnaires that are not held in copyright, or use them and relevant articles as resources to help clarify your own ideas about concepts without claiming the same level of reliability for results.

It is useful to do a literature search by Internet for appropriate standardized instruments and articles about their use. The American Psychological Association (APA) has a webpage of information for people looking for suitable standardized tests. See <http://www.apa.org/science/faq-findtests.html>. See also: TestLink at <http://www.ets.org/testcoll/index.html>), Test Locator from ERIC/AE at <http://www.ericfacility.net/ericdigests/ed385604.html> and the Health and Psychological Instruments (HAPI) database through a library site. Test Reviews are online at the Mental Measurements Yearbook at: <http://buros.unl.edu/buros/jsp/search.jsp>.

4.3 Look for model questionnaires

A variety of non-standardized questionnaires have been developed by agencies and community groups, which can provide models or be used intact with permission. Questionnaires from community groups are likely to have a mix of question types to collect quantitative and qualitative data.

Advantages

- already been tested and used
- development work is done regarding
 - how much to ask
 - question wording
 - placement
 - questionnaire format

Characteristics of good model questionnaires

- contain a description about confidentiality
- provide some explanation for the purpose and use of the questionnaire
- easy to read and follow from question to question
- reasonably consistent in format and layout of questions
- give precise, comprehensible instructions
- differentiate instructions clearly from questions (e.g., font size, type, boldface)
- express appreciation for the respondent's time and assistance
- have easily understood, unambiguous questions
- are appropriate for the purpose and the respondent in tone and content
- unemotional, impersonal and reasonably unintrusive
- provide opportunities for identifying date, collection place and coded identification as needed
- provide simple ways to minimize the amount of writing an informant needs to do, e.g., tick boxes, lists of responses to circle
- provide prompts to assist specific answers and avoid unclear responses (“yr/month/day” instead of “date”; birthdate instead of “age”)
- use scales to identify the intensity of the response (e.g., a 5-point scale with each point labeled)

Suitability

Even in good model questionnaires, some questions may not be suitable as they are worded.

- omit unnecessary or unsuitable questions unless the intention is to compare the results of the model questionnaire as a whole to your own
- make every effort to contact the originator of the questionnaire and obtain consent to use it

Determine the utility of data to be gathered from a question before including it in the final questionnaire.

4.4 Modify existing questions to suit respondents

Consider both the ability and willingness of respondents to answer the questions, in terms of, e.g.,

- language, reading level and specialized vocabulary
- knowledge
- cultural experience, attitudes
- community context

It is important that questions are interpreted the same way by all respondents, so they are all answering the same question—the one you meant.

Language and vocabulary

Translation

When language comprehension is a problem and a written questionnaire is necessary:

- translation of questions into the most familiar language for informants is a good option but may introduce new errors.
- a helpful check is to have a second person translate questions back into the first language (without seeing the original)
- compare results of re-translation and discuss implications of any differences.

Other options

Options followed by youth programs

- have an oral interview conducted by a staff person or volunteer who speaks the same language

- children often act as translators for parents in community programs
- someone else may be able to explain the questionnaire at the time it is being completed
- errors may creep in because there is no way of ensuring the explanations are sufficiently correct or complete

Inappropriate vocabulary

Inappropriate language may have a negative effect on respondents' willingness to answer a question

- possible problems in vocabulary include:
 - jargon or specialized meanings
 - convoluted phrases open to more than one interpretation
 - juvenile language that undermines the serious intent of the questions
 - culturally offensive terms

Generally a grade 4-6 vocabulary level is considered simple enough without seeming too juvenile. (Available software programs can determine vocabulary level.)

Knowledge

View each question from the viewpoint of the respondent, e.g., youth participating in the program, parent, teacher, volunteer or tutor.

- Can each question be answered easily, without much mulling over and second-guessing?
- Is it equally easy for all informants to answer the question?
- Does the respondent have the knowledge to answer the question?
- Does the question ask about a concept that needs a definition or some example?
- Does the question ask about something outside the informant's sphere of experience, e.g., asking parents about how children are behaving in the classroom or teachers about conditions or behaviour in a child's home?
- Is the question culturally appropriate or sensitive to cultural, religious or ethnic differences? For example, the terms synagogue, mosque or simply "religious services" might be included to expand a question about church attendance.
- Does the question ask for secondhand information?

Clarify whether a question should ask about attitudes or about behaviour and which is more likely to provide the most meaningful information. Examples of behavioural change can be a strong indicator of program effectiveness. Reporting on behaviour may also be less subjective or influenced by a respondent's support for the program. Questions about observed behaviour can be asked of several sources to build a composite picture and strengthen data.

Need an example?

For example, “How did your child/student enjoy the program?” really means,

“How do you think they enjoyed the program?” The reliability of the response could vary depending on the parent or teacher’s level of knowledge and relationship with the child. Firsthand information can be obtained by reworking the question, e.g., “Has your child/student ever expressed any feelings about the program? Yes/ no If yes, what did they say?”

Community context

Words may assume different meanings in different contexts, providing stumbling blocks to confuse data analysis. Specific terms may be unknown in a different jurisdiction. It is helpful to check terms with a similar audience to the informants and convert them to local usage. Obvious examples are U.S. compared to Canadian usage, in currency, terms such as food stamps vs. food banks or government programs such as the U.S. *National School Lunch Program*.

Questions should be unambiguous. Less obvious problems may be arise from questions that are inappropriate for the local context. Asking a general or open-ended question allows a fuller range of responses that can be classified later. Focus groups or more general discussions with respondents before writing a questionnaire also helps determine the full range of options to be considered.

Need an example?

For example, a researcher inquiring about relationships with non-school activities and violence asked youth in a particular area about membership in non-school clubs. Since there were few non-school clubs available in the area, the number of positive responses led the researcher to suspect that some youth had considered gangs to be non-school clubs! It can be useful to add examples. In the above case, adding, “e.g., Boy Scouts, 4-H, Boys and Girls clubs” might have given the researcher more confidence in the data. What if youth also have access to athletic clubs for golf, tennis, swimming, curling or gymnastics? Is membership in those clubs any different from membership in the YM/YWCA, which respondents would likely not consider a “club”? Beginning with a more general question, then asking for specifics is another way of avoiding ambiguity.

For example, “Do you belong to a non-school group or club where you spend time outside school? If yes, please specify what it is.” _____.

4.5 Guidelines for writing new questions:

Question types

Open-ended questions

- allows the informant to answer freely
- useful when you cannot predict all possible responses and want to know the range of available options
- more time-consuming for the respondent, who may then choose to skip them
- handwritten responses more time-consuming to record and analyse
- may be difficult to decipher or ambiguous

Closed-ended questions

- gives a set number of responses to choose from
- useful when the question writer knows what range of responses is available
- adding an “other” category with a “please specify”, takes care of unforeseen options but is more time-consuming to code.

Response options

Checklists

- include instructions to “**check ONLY one**”, or “**check all that apply**”.
- if the latter, any box left blank by accident would be coded as a “NO” response
- avoid that problem by providing options for every item, e.g, yes, no; never, sometimes, often, always

Scales

Since devising good scales is more difficult than it seems, try to find scales that have been used somewhere else previously.

- providing a scale for a response e.g., from 1-5 or from poor to excellent, is common

- options need to be mutually exclusive and not confusing
- options cover the full range of likely replies and are evenly spaced
- generally include an odd number of options, (often 5 or 7 points) to allow for a neutral mid-point
- omitting a neutral midpoint forces respondents to state an opinion by not giving them the “I don’t know” option
- a Likert Scale asks for an opinion about a statement, e.g, “Please read the statements below and then circle the one number on the right that best describes how much you agree or disagree with each statement.”
- keep response scales in the same order from question to question (1-5) not (5-1), e.g., “I feel comfortable with my volunteer tutor.” (1-5 from total disagreement to total agreement) and “I am afraid of being bullied by older children in the program.” (1-5 from total disagreement to total agreement) (When coding the latter question, keep the scoring in the same positive direction as the first question by reversing the numbers. See Step 6.)
- positive statements are usually preferable unless a negative will give more relevant data

Provide sufficient information to informants

- provide definitions of standard terms (e.g., wages, household income) on the questionnaire or
- have staff give definitions while distributing questionnaire
- avoid defining words related to opinions or feelings
 - **informants may feel inhibited or restricted to a particular range of responses**
- provide examples instead of definitions to suggest types of feelings or opinions

Assumptions

- review questions to look for hidden assumptions
- discuss the meaning with others to suggest other interpretations
- informants can misunderstand and misconstrue questions that writers think are obvious
- e.g., Does the question specify a clear time frame or frequency, e.g., “in the past 12 months”, “this school year?”

Neutral questions

Avoid questions that anticipate one option more than another or do not provide a full range of options. For example, asking a child, “How much did you like being in the program, a lot or a little?” provides no possible response for someone who didn’t like it at all. Giving more options, with a broader range, provides more precise, therefore more meaningful data.

One concept per question

- avoid multiple ideas in a question (often containing “and”), where a respondent can answer “yes” to one part of the question and “no” to another part
- divide such questions into two separate questions that can be coded and compared separately

Follow-up questions

- consider following up questions about attitudes or opinions by asking about the intensity with which they are held.
- link opinion to behaviour to help evaluate the strength of people’s responses

Need an example?

A question asking how much someone liked reading a certain book, might have a follow-up, “Did you like the book well enough to recommend it to someone else?” or even better, “Did you recommend it to someone else afterwards?” or “Have you looked for other books since then by the same author?”

Verifying responses

It is helpful to verify the consistency or accuracy of responses by asking more than one question about a certain attitude or behaviour in a different way, where space allows.

Encourage precision

- avoid grouping data or collapsing categories for essential questions since you may need more precise data later on.
- there is a trade-off between missing opportunities for analysis and asking for so much precision that informants can’t or won’t want to provide it
- pretesting can help determine the best balance for each questionnaire.

For example, it would likely be more helpful for data analysis to ask an open-ended question, “How many times did you attend the program this session?” instead of giving a range of options such as “<4, 5-9, > 10.”

Personal questions

One major exception to precision concerns certain personal data. Respondents may skip questions they find too personal and possibly not complete the questionnaire. There may be less intrusive ways of finding out the information, e.g., on the subject of income, providing a checklist of options with a fairly wide range (>\$5,000.) will improve response rate.

- Other information can stand in for income, such as educational level and home ownership, which are good indicators of income and less likely to bother respondents.

The 'not applicable' category

A full range of responses may include “not applicable.” Questions that cannot be answered because they do not apply to respondents can be irritating and possibly influence their attitude towards completing the questionnaire.

Skipped responses (non-responses) may have been skipped on purpose or by accident, creating ambiguity for analysis. Think about all possible interpretations when writing questions and as needed add, “Please explain.” or ask a filter question.

Need an example?

Consider the question, “Does your child enjoy reading at the library?” Adding an option “not applicable” would cover respondents whose children do not read at the library. But a respondent might also check it to mean “I don’t know”, “my child does not use a library” or even “my child does not know how to read.”

Filter questions

- a preliminary question to filter out respondents who lack some knowledge or experience necessary to answer a question
- e.g., “Do you like playing checkers?” implies that the respondent knows what checkers is and has played it. Ask respondents first, “Have you ever played checkers?” and have those who answer “NO” skip to another question.

Skip patterns

- instruction to skip a response after a question filters out some respondents
- can be confusing and distracting
- consider using arrows if there are several skips
- best placement is directly after the response, to prevent the informant from overlooking the instruction
- phrase the skip as a positive statement, e.g.,

1. Do you take dance lessons?
1 Yes ____
2 No ____ **go to question 3**
instead of "If NO, do not answer question 2"

Make forms easy to complete

- try for consistency in typeface, layout of similar questions
- more white space on forms, especially between questions, increases readability and is more inviting for respondents
- reduce the amount of writing respondents need to do, e.g., provide check boxes, circles or ()
- provide sufficient space for written answers
- create logical order and flow to questions
- provide instructions for each question or each different type of question
- make instructions more noticeable with a different font or boldface
- keep instructions and questions together, not on separate pages
- start each sentence on a new line for easy readability

Provide information about confidentiality (also See Step 5)

Explain the purpose for collecting data on the form or in a verbal introduction (preferably both). Tell informants about confidentiality: that their responses will not be reported individually or with their names. It may be necessary to reassure participants that the program welcomes both positive and critical comments in order to identify gaps in service and strengthen programming. If personal questions are being asked, the need for confidentiality, and confidence that it will be maintained, are especially important. Some projects may require more safeguards around confidentiality than others. Since programs that want to follow individuals over time within a program or over various sessions, need to be able to identify the

participants who provided data, it is important to have some identifier, either name or another unique identifier like gender coupled with date of birth. Check that the identifier is unique. Data may be kept anonymous to all but one person who tags each questionnaire with an identification code instead of a name, records it on a master list, then links it to other data records by the same code. The master list should be kept secure and accessible to only one person on the evaluation team.

Even if names are collected, information provided in reports should not be traceable to an individual either by name or by specific identifying information. A personal statement such as a testimonial or quotation could be published with permission from the individual (and/or parents). However, families and programs may not anticipate all the possible negative consequences of releasing personal information that is distributed publicly.

4.6 Pre-test questionnaires

- a 'practice run' using a few people similar to informants in language, culture and education
- before pre-test, practise coding and recording data by preparing a dummy table of information to ensure neither gaps nor unnecessary questions. (See Step 5)
- watch for any hesitation and ask about the respondent's reaction
- after pre-test, discuss the informant's interpretation of questions to resolve possible misunderstanding

