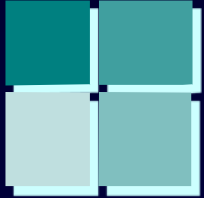
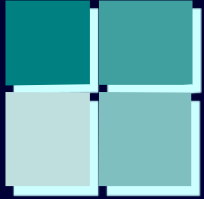


Selecting and Constructing Data Collection Instruments



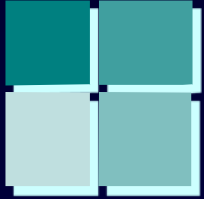
Introduction

- Data Collection Strategies
- Characteristics of Good Measures
- Quantitative and Qualitative Data
- Tools for Collecting Data



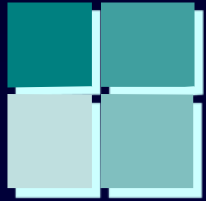
Data and Information

- Most people use the terms data and information interchangeably. But data and information are not the same thing! The distinction is important when selecting and constructing data collections instruments.



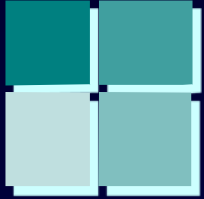
Data

- Data are raw facts in isolation. These isolation facts convey meaning but generally are not useful by themselves. Data consists of factual elements that describes some object or event. Data can be thought of as raw numbers or text. The five major types of data are *numbers* , *text*, *images*, *sound and video* to handle more sophisticated data and more difficult tasks, computer hardware and software have grown increasingly complex.



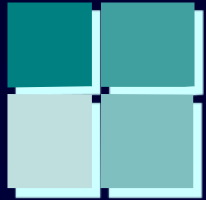
Information

- Information is data that has been manipulated to be useful to someone.
- An organized meaningful and useful interpretation of data (raw facts, figures & details).
- An information system creates , collects, and stores data and processes that data into useful information. So, information is a function of data and processing.



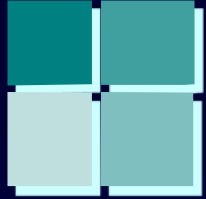
Data Collection Strategies

- No best way: decision depends on:
 - What you need to know: *numbers or stories*
 - Where the data reside: *environment, files, people*
 - Resources and time available
 - Complexity of the data to be collected
 - Frequency of data collection
 - Intended forms of data analysis



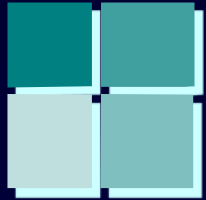
Rules for Collecting Data

- Use multiple data collection methods
- Use available data, but need to know
 - how the measures were defined
 - how the data were collected and cleaned
 - the extent of missing data
 - how accuracy of the data was ensured



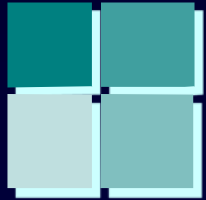
Rules for Collecting Data

- If must collect original data:
 - be sensitive to burden on others
 - pre-test, pre-test, pre-test
 - establish procedures and follow them (protocol)
 - maintain accurate records of definitions and coding
 - verify accuracy of coding, data input



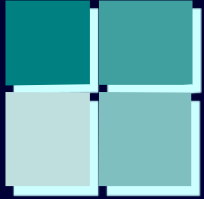
Structured Approach

- All data collected in the same way
- Especially important for multi-site and cluster evaluations so you can compare
- Important when you need to make comparisons with alternate interventions



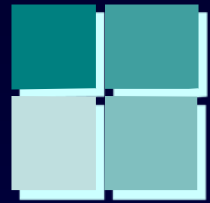
Use Structured Approach When:

- need to address extent questions
- have a large sample or population
- know what needs to be measured
- need to show results numerically
- need to make comparisons across different sites or interventions



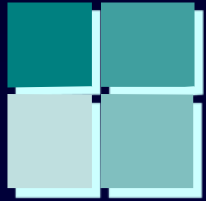
Semi-structured Approach

- Systematic and follow general procedures but data are not collected in exactly the same way every time
- More open and fluid
- Does not follow a rigid script
 - may ask for more detail
 - people can tell what they want in their own way



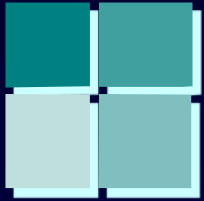
Use Semi-structured Approach when:

- conducting exploratory work
- seeking understanding, themes, and/or issues
- need narratives or stories
- want in-depth, rich, “backstage” information
- seek to understand results of data that are unexpected



Characteristics of Good Measures

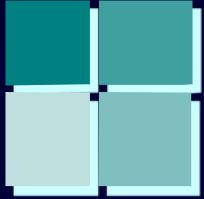
- Is the measure relevant?
- Is the measure credible?
- Is the measure valid?
- Is the measure reliable?



Relevance

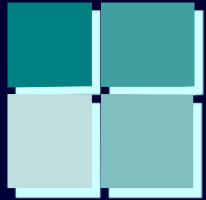
Does the
measure capture
what matters?

Do not measure
what is easy
instead of what is
needed



Credibility

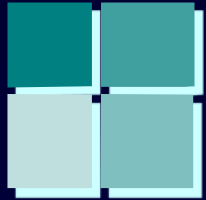
Is the measure believable? Will it be viewed as a reasonable and appropriate way to capture the information sought?



Internal Validity

How well does the measure capture what it is supposed to?

Are waiting lists a valid measure of demand?

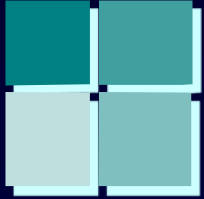


Reliability

A measure's precision and stability- extent to which the same result would be obtained with repeated trials

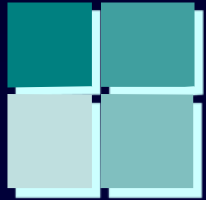
How reliable are:

- birth weights of newborn infants?
- speeds measured by a stopwatch?



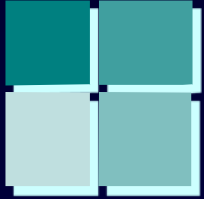
Quantitative Approach

- Data in numerical form
- Data that can be precisely measured
 - age, cost, length, height, area, volume, weight, speed, time, and temperature
- Harder to develop
- Easier to analyze



Qualitative Approach

- Data that deal with description
- Data that can be observed or self-reported, but not always precisely measured
- Less structured, easier to develop
- Can provide “rich data” – detailed and widely applicable
- Is challenging to analyze
- Is labor intensive to collect
- Usually generates longer reports



Which Data?

If you:

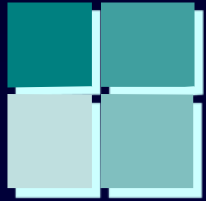
- want to conduct statistical analysis
 - want to be precise
 - know what you want to measure
 - want to cover a large group
-

- want narrative or in-depth information
- are not sure what you are able to measure
- do not need to quantify the results

Then Use:

Quantitative

Qualitative



Obtrusive vs. Unobtrusive Methods

Obtrusive

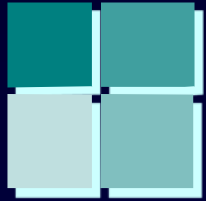
data collection methods that directly obtain information from those being evaluated

e.g. interviews, surveys, focus groups

Unobtrusive

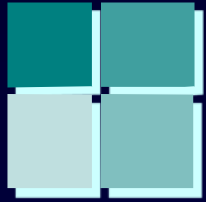
data collection methods that do not collect information directly from evaluated-

e.g., document analysis, Google Earth, observation at a distance.



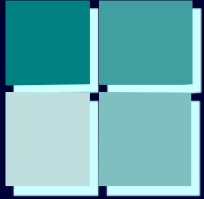
How to Decide on Data Collection Approach

- Choice depends on the situation
- Each technique is more appropriate in some situations than others
- Caution: All techniques are subject to bias



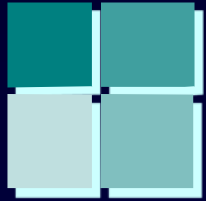
Triangulation to Increase Accuracy of Data

- Triangulation of methods
 - collection of same information using different methods
- Triangulation of sources
 - collection of same information from a variety of sources
- Triangulation of evaluators
 - collection of same information from more than one evaluator



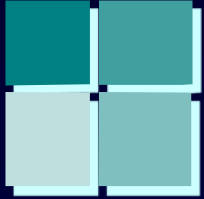
Data Collection Tools

- Participatory Methods
- Records and Secondary Data
- Observation
- Surveys and Interviews
- Focus Groups
- Diaries, Journals, Self-reported Checklists
- Expert Judgment



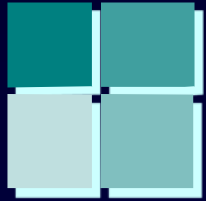
Tool 1: Participatory Methods

- Involve groups or communities heavily in data collection
- Examples:
 - community meetings



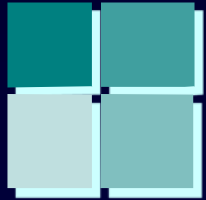
Community Meetings

- One of the most common participatory methods
- Must be well organized
 - agree on purpose
 - establish ground rules
 - who will speak
 - time allotted for speakers
 - format for questions and answers



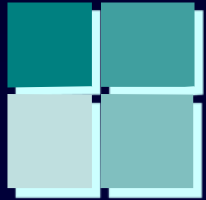
Tool 2: Records and Secondary Data

- Examples of sources:
 - files/records
 - computer data bases
 - industry or government reports
 - other reports or prior evaluations
 - census data and household survey data
 - electronic mailing lists and discussion groups
 - documents (budgets, organizational charts, policies and procedures, maps, monitoring reports)
 - newspapers and television reports



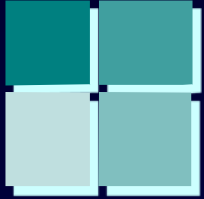
Using Existing Data Sets

Key issues: validity, reliability, accuracy, response rates, data dictionaries, and missing data rates



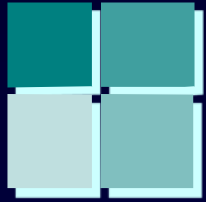
Advantage/Challenge: Available Data

Advantages	Often less expensive and faster than collecting the original data again
Challenges	There may be coding errors or other problems. Data may not be exactly what is needed. You may have difficulty getting access. You have to verify validity and reliability of data



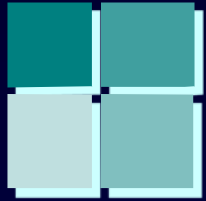
Tool 3: Observation

- See what is happening
 - traffic patterns
 - land use patterns
 - layout of city and rural areas
 - quality of housing
 - condition of roads
 - conditions of buildings
 - who goes to a health clinic



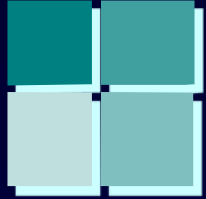
Observation is Helpful when:

- need direct information
- trying to understand ongoing behavior
- there is physical evidence, products, or outputs that can be observed
- need to provide alternative when other data collection is infeasible or inappropriate



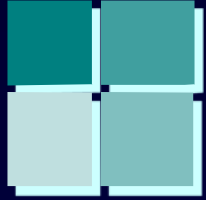
Degree of Structure of Observations

- Structured: determine, before the observation, precisely what will be observed before the observation
- Unstructured: select the method depending upon the situation with no pre-conceived ideas or a plan on what to observe
- Semi-structured: a general idea of what to observe but no specific plan



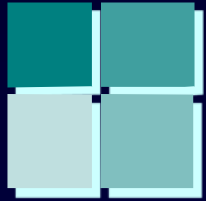
Google Earth

- Maps and satellite images for complex or pinpointed regional searches
- Has an Advanced version and an Earth Outreach version
- Web site for Google Earth
 - <http://earth.google.com/>



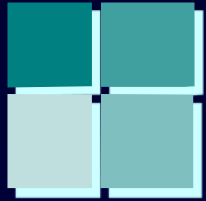
Ways to Record Information from Observations

- Observation guide
 - printed form with space to record
- Recording sheet or checklist
 - Yes/no options; tallies, rating scales
- Field notes
 - least structured, recorded in narrative, descriptive style



Guidelines for Planning Observations

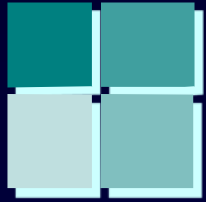
- Have more than one observer, if feasible
- Train observers so they observe the same things
- Pilot test the observation data collection instrument
- For less structured approach, have a few key questions in mind



Advantages and Challenges: Observation

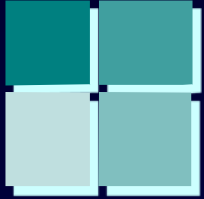
Advantages Collects data on actual vs. self-reported behavior or perceptions. It is real-time vs. retrospective

Challenges Observer bias, potentially unreliable; interpretation and coding challenges; sampling can be a problem; can be labor intensive; low response rates



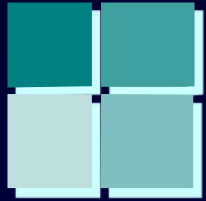
Tool 4: Surveys and Interviews

- Excellent for asking people about:
 - perceptions, opinions, ideas
- Less accurate for measuring behavior
- Sample should be representative of the whole
- Big problem with response rates



Structures for Surveys

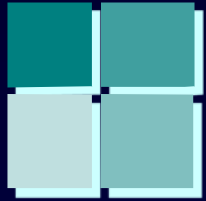
- **Structured:**
 - Precisely worded with a range of pre-determined responses that the respondent can select
 - Everyone asked exactly the same questions in exactly the same way, given exactly the same choices
- **Semi-structured**
 - Asks same general set of questions but answers to the questions are predominantly open-ended



Structured vs. Semi-structured Surveys

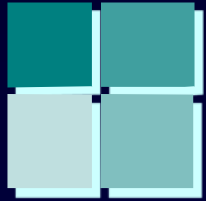
Structured harder to develop
 easier to complete
 easier to analyze
 more efficient when working with large numbers

Semi-structured easier to develop: open ended questions
 more difficult to complete: burdensome for people to complete as a self-administrated questionnaire
 harder to analyze but provide a richer source of data, interpretation of open-ended responses subject to bias



Modes of Survey Administration

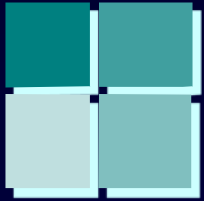
- Telephone surveys
- Self-administered questionnaires distributed by mail, e-mail, or websites
- Administered questionnaires, common in the development context
- In development context, often issues of language and translation



Advantages and Challenges of Surveys

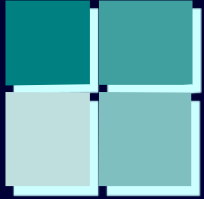
Advantages Best when you want to know what people think, believe, or perceive, only they can tell you that

Challenges People may not accurately recall their behavior or may be reluctant to reveal their behavior if it is illegal or stigmatized. What people *think they do* or *say they do* is not always the same as what they *actually do*.



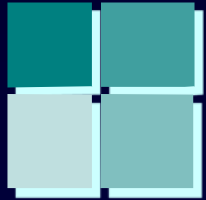
Interviews

- Often semi-structured
- Used to explore complex issues in depth
- Forgiving of mistakes: unclear questions can be clarified during the interview and changed for subsequent interviews
- Can provide evaluators with an intuitive sense of the situation



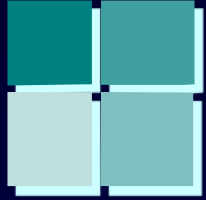
Challenges of Interviews

- Can be expensive, labor intensive, and time consuming
- Selective hearing on the part of the interviewer may miss information that does not conform to pre-existing beliefs
- Cultural sensitivity: e.g., gender issues



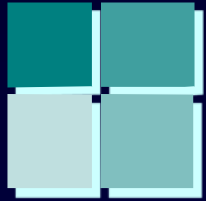
Tool 5: Focus Groups

- Type of qualitative research where small homogenous groups of people are brought together to informally discuss specific topics under the guidance of a moderator
- Purpose: to identify issues and themes, not just interesting information, and not “counts”



Focus Groups Are Inappropriate when:

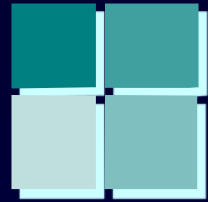
- language barriers are insurmountable
- evaluator has little control over the situation
- trust cannot be established
- free expression cannot be ensured
- confidentiality cannot be assured



Advantages and Challenges of Focus Groups

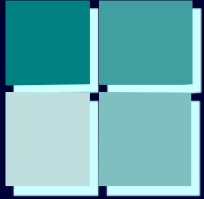
Advantages Can be conducted relatively quickly and easily; may take less staff time than in-depth, in-person interviews; allow flexibility to make changes in process and questions; can explore different perspectives; can be fun

Challenges Analysis is time consuming; participants not be representative of population, possibly biasing the data; group may be influenced by moderator or dominant group members



Tool 6: Diaries and Self-Reported Checklists

- Use when you want to capture information about events in people's daily lives
- Participants capture experiences in real-time not later in a questionnaire
- Used to supplement other data collection



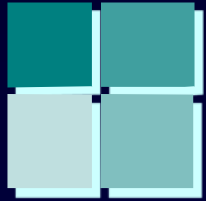
Advantages and Challenges of Diaries and Self-reported Checklists

Advantages

- Can capture in-depth, detailed data that might be otherwise forgotten
- Can collect data on how people use their time
- Can collect sensitive information
- Supplements interviews provide richer data

Challenges

- Requires some literacy
- May change behavior
- Require commitment and self-discipline
- Data may be incomplete or inaccurate
- Poor handwriting, difficult to understand phrases



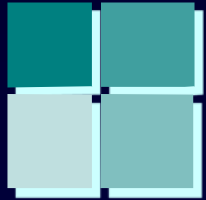
Tool 7: Expert Judgment

Use of experts, one-on-one or as a panel

E.g., Government task forces, Advisory Groups

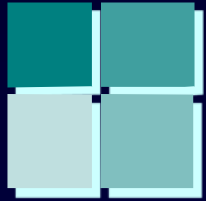
Can be structured or unstructured

Issues in selecting experts



Selecting Experts

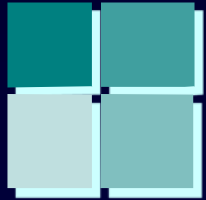
- Establish criteria for selecting experts not only on recognition as expert but also based on:
 - areas of expertise
 - diverse perspectives
 - diverse political views
 - diverse technical expertise



Advantages and Challenges of Expert Judgment

Advantages Fast, relatively inexpensive

Challenges Weak for impact evaluation
May be based mostly on perceptions
Value of data depends on how credible the experts are perceived to be

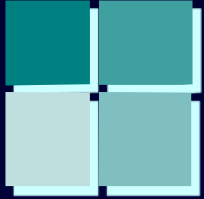


Data Collection Summary

Choose more than one data collection technique

No “best” tool

Do not let the tool drive your work but rather choose the right tool to address the evaluation question



A Final Note....

“I never guess. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts and theories, instead of theories to suit facts.”

--Sir Arthur Conan Doyle



Questions?