

Data Collection Primer

Evaluation models often explain what needs to be done but rarely how to do it. This chapter presents a primer on data collection methods. The precise way these activities are performed depends on the evaluation context and the information needed. However, the basic principles apply regardless of the context.

Data Quality

The first and most important principle related to data collection is quality. A common way to remember this principle is GIGO (Garbage In, Garbage Out). No matter how well you design your evaluations, if the information you collect is inaccurate or incomplete, the interpretations of your results and the recommendations you make will be flawed. One of the biggest mistakes researchers and evaluators make is assuming the data they collect or are provided is accurate without taking steps to confirm the data is accurate (i.e., trust but verify, triangulate).

Data collection can go wrong in many different ways, depending on the purpose for collecting data and the methods being utilized. When the information we collect is incorrect, we call this measurement error, bias, or noise. Sometimes the data errors result from inadequate data collection instruments; other times, bias occurs because the respondent is unable or unwilling to provide the information we wish to obtain. Errors can also occur when interpreting the data we obtain. Measurement error is described differently and is subject to different kinds of errors depending on the type of data you are collecting.

Self-report Data Collection Issues. The term *response bias* is commonly used to describe issues that occur when collecting self-report data (i.e., surveys and interviews). Response bias occurs in many ways. These errors occur because the respondent cannot or will not provide accurate data. The accuracy of the data we collect affects the integrity of the results and the credibility of our interpretations. Self-report data collection is most appropriate when evaluating user satisfaction, usability, and UX testing (i.e., learner reactions, perspective, and feeling). Several types of response bias associated with surveys and interviews include: (see [source](#))

Recall Bias. This is common in self-report situations when respondents are asked to provide information retrospectively. Human memory is imperfect. Some information is more likely to be remembered than others. A person's ability to recall events and feelings will depend on the metacognitive ability of the individual and the significance of the event to that particular person (i.e., a *vividness effect*). Recall often depends on the time interval between the event and when the individual is asked to recall their perceptions. A person may have forgotten the event altogether; they may remember incorrectly or revise their recollection (see prestige bias).

Social Desirability & Conformity Bias. It can be hard for respondents to openly express non-conformity when asked to self-report their behavior, beliefs, and opinions; this is especially true when the respondent believes they may be ridiculed or despised. In such cases, respondents tend to provide a socially acceptable response (sometimes subconsciously) over their true feelings. For example, a respondent may tend to agree with a statement more strongly than how they truly feel when the item addresses something that is generally seen in society as desirable or expected.

Prestige Bias or Report Bias. This bias is related to social desirability bias but it is based on an individual's personal desire to be seen in a positive light. This bias is based on personal feelings, not a general instinct for conformity. It involves selectively revealing or suppressing information. For example, respondents may round up

instead, the individual may honestly remember the facts inaccurately (i.e., they have revised details of the event in their mind). They believe what they remember as true, even though it is not. Respondents often tend to view or recall their own situation in a more favorable light than is actually the case—subconsciously protecting their self-image or inflating their ego. It is often good practice to assume that if a question has a potential prestige component, the responses are likely inflated to present the respondent more favorable. Exactly how much they are inflated will depend on the question, context, and respondents.

Acquiescence or Agreement Bias. This bias is like conformity bias. However, unlike conformity bias, in this case, the respondent will, in general, and inadvertently, agree with statements. With this bias, participants tend to select a positively worded response option (i.e., a *framing effect*) or disproportionately indicate a positive direction to their feelings. This bias will skew results towards the positive.

Item and Option Order Effect Bias. Order bias can be the result of both item order and response option order. The order in which survey items are presented can affect a respondent's answers due to a *priming effect*. People tend to contextualize their responses. Because of this, survey or interview questions that come just before a particular query may provide information that respondents will use as a context when formulating their subsequent answers. This is not always a negative. It is only important to know that if an alternate primer was presented, the responses might be significantly different.

Two common response biases associated with response option order are primacy and recency bias. *Primacy bias* is the tendency for respondents to pick one of the first viable options presented to them. This can happen when a respondent quickly reads through the survey and picks one of the first response options they agree with. *Recency bias* is the tendency to pick an answer option presented at the end of a list. This is especially problematic when a long list of options is presented; the choices respondents read last are more memorable, so they tend to select answers near the end of the list.

Mood Bias and Emotional Mind-Sets. One's mood or mindset will affect the way responses are provided. For example, if a participant is exceptionally happy or angry for some reason while taking a survey, their emotional state affects the general pattern of responses provided. Given time, the respondent's current extreme emotions may subside, which will modify the intensity of the responses provided. Emotional responses can be intense in either a positive or negative direction. You will also see this when the survey addresses an emotionally charged topic. Responses may tend to be on the extreme ends of the response scale, possibly because those who choose to complete the survey have strong opinions; however, mood bias becomes a problem when the respondent's current emotional state temporarily exaggerates opinions.

Central Tendency Bias. This bias refers to the tendency of some individuals to avoid responding in extreme ways. For example, some people may never indicate they strongly agree or are extremely dissatisfied (i.e., nothing is perfect, and nothing is entirely without merit). This is the opposite of a mood bias in that responses from those who have this bias will trend closer to the center of the response scale.

Demand Characteristic Bias. A demand characteristic is used to describe specific cues in research that may inadvertently influence a participant's response. A demand characteristic can manifest in a number of different ways if the researcher is not careful when designing and proceeding with a study. In social science research, demand characteristics can create bias when the subject becomes aware of the purpose of the study (i.e., a *Hawthorn effect*). This may potentially bias or invalidate the outcomes. When a respondent becomes aware of the reason or purpose of the study, they may intentionally provide answers they feel would influence the results. For example, if a respondent figured out that the results of a survey will be used to set policy, the individual may attempt to answer in a way that they feel would be beneficial to them.

Random Response Bias. Random response bias can occur when a respondent honestly does not know the answer to the question but answers anyway. This can happen when you ask a respondent to answer a question

another individual, prompting a random response bias.

Another way this bias can manifest is when an individual has an opinion but hasn't considered their true feelings carefully. Like a central tendency bias, these individuals also tend to choose options toward the middle of the response scale. At times, people with this bias will choose the exact middle point (on an odd-numbered response scale) simply because they don't want to think about the issue or don't really care. This bias can also manifest itself maliciously when an individual intentionally responds in a random fashion without actually reading the survey questions or carefully considering what is being asked.

Additional issues can affect the quality of data obtained through interviews. There are several ways to conduct an interview, including structured interviews, unstructured interviews, and focus groups. In addition to response bias issues, interviews are also susceptible to problems regarding how the interviewer conducts the interview (see [moderated user testing](#)). A few examples include:

Authority Bias. This happens when an evaluator values information from what they perceive as a reputable source—disbelieving or devaluing other sources of information.

Personality Issues. This is a concern when gathering data using a focus group. A focus group can be more efficient than individual interviews but only if they are conducted properly. The makeup of the group is also essential. A focus group can be a synergetic dialogue or a one-sided conversation. Understanding personality characteristics matters because your goal is to hear the voices of all participants and understand the issues fully from various perspectives. People who are shy, insecure, non-confrontational, or prone to authority bias are less likely to share their thought and feeling. In contrast, individuals who are outgoing, assured, opinionated, angry, or in a position of power over others in the group may take over a conversation or restrict the free flow of honest dialogue.

Measurement Issues. Data collection quality issues also occur when using tests that measure cognitive and affective learning outcomes. The terms *valid* and *reliable* describe the quality of data obtained in this manner. Validity in assessment refers to the degree to which a test measures what it was designed to measure, and reliability refers to the consistency of the results. A measurement instrument must adequately and consistently provide evidence (i.e., results) of whether the learning objectives associated with an instructional product have been achieved. An assessment instrument may provide a measure of student learning and ability, or a measure of an individual's attitude, personality, or beliefs (e.g., a [scale](#)). In either case, the measurement instrument must adequately align with and target all aspects of the intended learning objectives. This type of data collection is most appropriate when conducting an effectiveness evaluation. An assessment instrument may be included in the design of the instructional products. However, when there are no specific learning objectives provided or the assessment tool is missing or inadequate, the evaluator may need to create a data collection instrument.

Another issue associated with assessment is that of interpretation resulting and the incorrect use of statistical methods. Method errors happen when the evaluator uses an inappropriate statistical procedure or fails to verify that assumptions have been met. Method errors occur when the evaluator is unskilled in interpreting the result correctly. Measurement issues can also happen when the results of a particular analysis or the assessment results are difficult to understand.

Observations Issues. Observations can be a valuable tool when collecting evaluation data. This type of data collection is used when conducting effectiveness evaluations, usability tests, implementation fidelity studies, and negative case analyses. In an effectiveness evaluation, observations are needed when the intended learning objectives involve performance. To properly assess performance, a rubric is needed that outlines the criteria you will use to judge whether the performance meets a specific standard. In other types of evaluation (e.g., usability testing), observations are needed to understand the phenomenon better. Observations can be completed in a laboratory or an authentic setting. Observing individuals in an authentic setting has the advantage of avoiding certain kinds of bias; however, they have

may exhibit some degree of performance anxiety. These situations are the result of *observer influence*.

In addition to the issue of observer influence, observations are particularly prone to interpretation error resulting from observer biases. These errors happen when the evaluator is unskilled or unaware of how a specific bias might influence them. Interpretation error occurs when the evaluator fails to consider alternative explanations, neglects to conduct member checks and peer reviews, or has a specific cognitive bias. Observer bias can also occur when the evaluator is unduly influenced by their personal values, the desire to obtain a specific result, or their inability to understand what they observed. These are often unintentional but not always. A few examples include:

Confirmation Bias. Confirmation bias describes the tendency to look for evidence that supports one's prior beliefs, ignore contradictory evidence, and interpret ambiguous information in a way that confirms a desired position or finding.

Belief Bias. Similar to confirmation bias, belief bias occurs when an evaluator judges the strength of an argument or the importance of evidence based on how well the evidence supports their values, beliefs, or previous understanding.

Attention bias (or blindness). This happens when an observer pays attention to specific details while ignoring or failing to see other potentially significant evidence. This often happens in structured observations when the observer is tasked with collecting specific data. This can also be the result of priming. When an observer is told to expect something, they tend to see it.

Attribution Bias. This is an interpretation error. It happens when the observer notes a specific action, behavior, or event and misunderstands its cause. For example, when you see someone smile, you might attribute the expression to the individual feeling happy, amused, love, or satisfied. However, alternative explanations might include feelings of superiority, disdain, disregard, fear, nervousness, or submission. We might also misinterpret the target of the emotion. We may assume the person is smiling at us when they might be looking at or thinking about something completely different. Likewise, observations are not suitable for identifying another person's intentions.

Halo Effect. A halo effect happens when an evaluator forms a positive or negative impression based on previous knowledge or experience. In these situations, the empirical evidence is ignored or excused.

Reporting Bias. This is typically applied to a respondent who selectively reveals or suppresses information. However, it can also apply to observers when they under or over report specific observations. It may happen when a person is hesitant to report something sensitive or potentially controversial. It can also happen when they over-emphasize observations they think are interesting or unique.

Cultural Bias. This bias involves the tendency for evaluators to interpret observations and judge them based on the values and standards of their own culture.

Data Collection Instruments

A second general principle related to data collection is asking the right questions in the right way.

Too often, evaluators fail to validate their data collection instruments and protocols adequately. Instruments can be flawed in several ways. They may fail to ask the right questions or the question may be asked in a way respondents cannot understand or misinterpret what is being asked. The questions that are posed may be well designed, but the instrument (or protocol) neglects to ask other important questions (a form of attention bias). An interviewer may also fail to prompt a participant to fully explain their answers, thinking they understood fully what was being said (an attribution issue). Pilot testing the instrument with someone from the target population can alleviate issues. Conducting member checks and peer reviews can help when interpreting results. (see [survey planning](#) example)

Another important principle of data collection is getting information from direct sources.

Basically, you need to get information from those who have the data—preferable an objective source. For example, you might want to know how often students attended a class or how much time they spent studying. You plan to use this information to disaggregate results as part of an implementation fidelity study or an effectiveness study to verify the instructional product is being used. When planning your data collection you need to consider where you will get the information you need. In this example, you might ask a student's parent how often their child attends class, but they may not be able to provide factual information – only their perceptions or beliefs. The student may be asked to recall how often they study but may be unwilling to respond accurately or may have forgotten. You might obtain the data from a homework log where students record when they started and ended their studies. This would be good if the students faithfully and honestly recorded their study time. Suppose the learning activities are contained in a technology-enabled learning management system (LMS). In that case, you might also rely on data analytics obtained from the system that recorded when a student accessed the online activities. However, even in this situation, you may encounter measurement errors. For example, a student may have logged into the LMS and accessed a homework assignment but also watched television, listened to music, took bathroom breaks, texted friends, and perused the refrigerator for sustenance. When planning your data collection, you need to make sure the source of information will provide accurate data. One of the biggest challenges can be getting access to those who have the information you need.

Chapter Summary

- Data collection is crucial in the evaluation process.
- Unfortunately, we often mistakenly assume that the data we collect is accurate and complete.
- The first principle of data collection is data quality. It is essential that we take steps to ensure our data is accurate and unbiased. We likely will never eliminate all measurement errors, but we can take steps to alleviate the likelihood that they will occur.
- The second principle of data collection is associated with the data collection instruments and protocols we use to collect information. Asking the right questions, in the right way is essential to obtaining the data we need to answer our evaluation questions.
- The last principle requires we use direct sources. We must obtain the information we need from those who have direct knowledge or direct access to the data we need.

Discussion Questions

1. Consider the data you plan to obtain for one of your projects. Describe which threats are most likely to affect the quality of your data? What can you do to alleviate these issues?

This content is provided to you freely by EdTech Books.

Access it online or download it at https://edtechbooks.org/eval_and_design/data_collection.