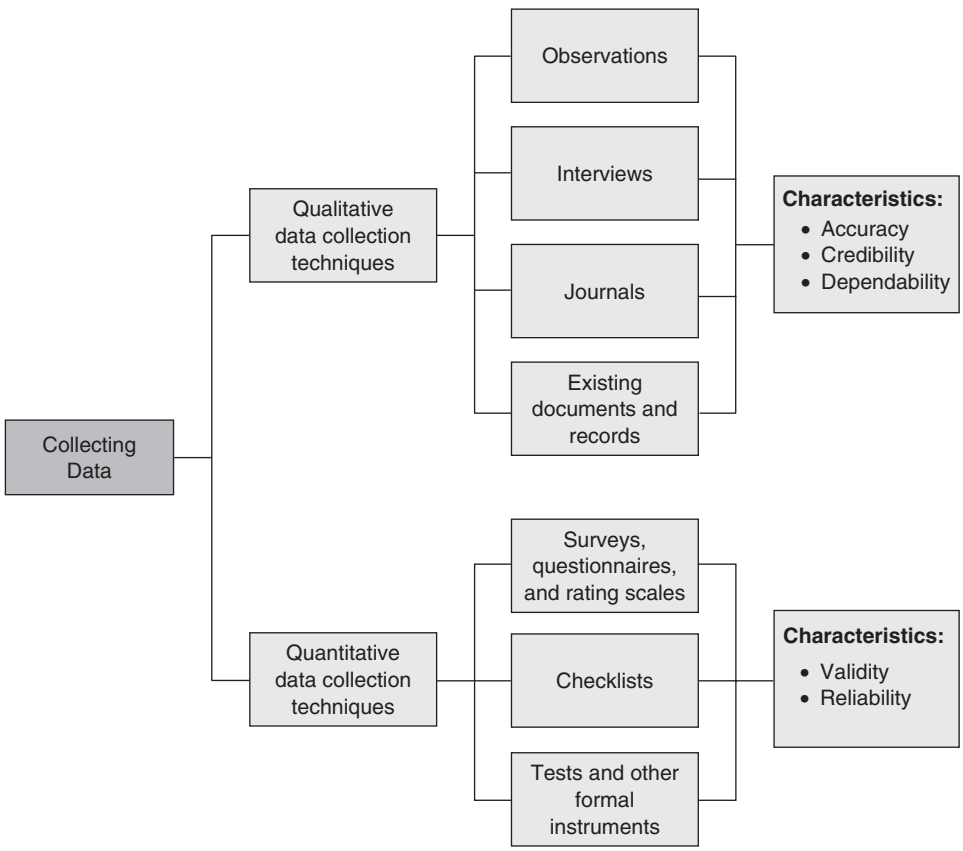
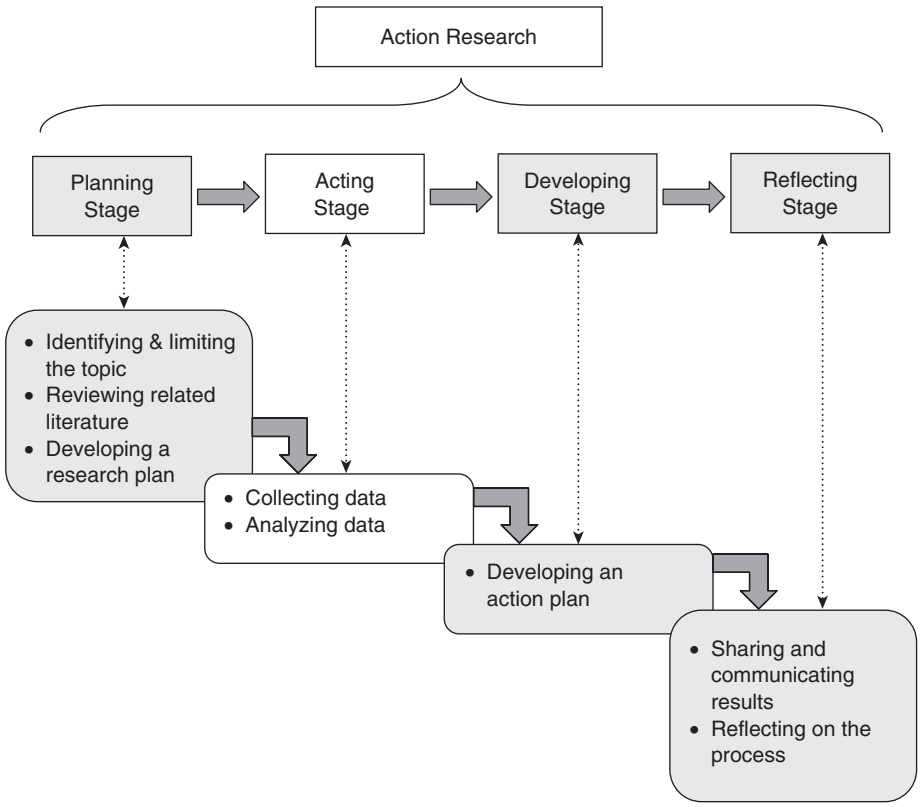
Chapter 5

COLLECTING DATA



In the this chapter, we enter into the second stage—the *acting* stage—of conducting a classroom-based action research project. Recall that the acting stage is comprised of data collection, which will be discussed in the present chapter, and data analysis, the topic for Chapter 6. As you will soon learn, there are numerous techniques that can be used to collect both qualitative and quantitative data for your teacher-led action research studies.



QUALITATIVE DATA COLLECTION TECHNIQUES

Recall that qualitative data is *narrative;* in other words, the data themselves are words. These "words" may appear in the form of interview transcripts, observational notes, journal entries, transcriptions of audio- or videotapes, or as existing documents, records, or reports. They may be collected using a variety of techniques, but it is important to remember that the resulting qualitative data will always consist of descriptive, narrative accounts.

Observations

As human beings, we are constantly observing and taking note of the world around us. Furthermore, as teachers, we are constantly observing our students. However, on a daily basis,

we typically observe our surroundings in somewhat of a haphazard manner—something more akin to "watching" than observing. **Observations**, as a means of collecting qualitative data, involve *carefully* watching and *systematically* recording what you see and hear going on in a particular setting (Schmuck, 1997). Observations can be extremely useful in certain situations where other forms of data collection simply will not work, such as when teachers want to check for students' nonverbal reactions to something that is occurring in the classroom or when students are working in small groups in order to better understand how they interact and communicate with one another.

Classroom observations can range from highly structured to semistructured to unstructured (Parsons & Brown, 2002; Schmuck, 1997). **Structured observations** typically require the observer to do nothing else but observe, looking usually for specific behaviors, reactions, or interactions. Because so many other things are going on in a given classroom when observations are being made, it is often difficult to conduct structured observations. Classroom-based action research should never be done at the expense of your teaching (Hubbard & Power, 2003; Johnson, 2005); it should only be done in order to enhance and inform your teaching. **Unstructured** or **semistructured observations** allow the teacher-researcher the flexibility to attend to other events or activities occurring simultaneously in the classroom or to engage in brief, but intense, periods of observation and note taking (Hubbard & Power, 2003). In addition, unstructured observations are more typical of qualitative data collection, since they are "free flowing," allowing the teacher-researcher to shift focus from one event to another as new, and perhaps more interesting, events arise (Leedy & Ormrod, 2005).

Schmuck (1997) discusses several advantages of conducting classroom observations. First, teachers can gather data about *actual* student behaviors, as opposed to asking students to report their perceptions or feelings. Second, this enables the teacher to see some things that students might not be able to report on themselves. Finally, as you will see shortly, devices such as videotape recorders allow teachers to "observe" even more than they would normally be able to with their own eyes.

However, conducting observations also has its limitations (Schmuck, 1997). First, the simple presence of the teacher as a "data collector"—with notebook and pencil, or perhaps a video recorder, in hand—can change student behavior. There is a great potential for them to behave differently or to say different things if they know that they are being watched carefully. Second, in cases where specific behaviors are sought, and since behavior can be adversely affected due to the presence of an observer, the teacher-researcher may have to wait for extended periods of time in order to observe the desired behavior. Still, the desired behavior may never occur, even if it is a normal, everyday occurrence. Finally, if teachers are working together on an action research project, different observers may see different things, even while observing the same event.

Classroom observations are usually recorded in the form of fieldnotes. Fieldnotes are written observations of what you see taking place in your classroom (Johnson, 2005). It can sometimes be overwhelming to try to record everything that you see, especially when trying to determine what is important (and, therefore, worth recording) and what is not. Johnson (2005) advises teacher-researchers to simply "stop thinking and just write what you see" (p. 63). As you observe and record what you see, you will undoubtedly begin to focus on things that are interesting or important. As you make observations over time, patterns will begin to emerge from the data you have collected.

When recording fieldnotes, you may want to consider dividing each page of your notebook into two columns. You should use the left column for recording your actual observations and

the right column for noting preliminary interpretations of what has been observed (Leedy & Ormrod, 2005). Bogdan & Biklen (1998) refer to these interpretations as **observer's comments**, or *OCs*. Observer's comments often shed light on the emerging patterns from your observational data. Including observer's comments in your observation notes is also one way to integrate reflection into the process of action research. The separation of these two types of commentaries is critical so that *actual* observations are not confused with what you think the observed event *means.* Teachers conducting action research studies need to remain as objective as possible in the records kept and data collected. As an aside, this need for objectivity also dictates that you not censure what you record in your notes with your "teacher's eyes"—do not hesitate to record something even if it reflects negatively on your teaching (Hubbard & Power, 2003); after all, you are trying to learn about and improve your professional practice. In addition, interpretations of observations may change over time as you collect more data; having a record of these changing interpretations can be invaluable over the course of your study. An example of a page from a book of fieldnotes I recorded several years ago during a study of positive reinforcement in a preschool setting, depicting this two-column format of actual observations and associated observer's comments, is shown in Figure 5.1.

Written fieldnotes can become problematic, however. They are often insufficient to depict the richness and the details of what one is observing (Leedy & Ormrod, 2005). **Videotapes** can provide assistance as a tool for recording observations, although they are not without their respective limitations, as well. Background noises may prevent you from hearing that on which you were hoping to focus your videotaped observation. Furthermore, video cameras can only capture what is happening in a given direction (i.e., the direction the camera is facing). Leedy & Ormrod (2005) suggest that, prior to beginning any formal observations, researchers should experiment and become familiar with various methods of recording observations in order to find what works best for the particular setting and situation. It is, however, important to remember that whatever mechanism you use to record your observations, you simply cannot physically record everything that you see or that is happening (Mills, 2003); it is best not to put pressure on yourself to try to do so.

On a practical note, several tips may facilitate your observations and the development of your observation skills. If you decide to observe and to record those observations using fieldnotes, you may want to consider carrying a clipboard or legal pad with you for several days prior to beginning your observations and recording any fieldnotes. It is important that the act of recording fieldnotes becomes a part of your daily routine, as opposed to something that "feels" unfamiliar, extraneous, or irrelevant. Similarly, if you decide that you will record your observations through the use of a video camera, you may want to set up the camera several days in advance of your recording. This is important because both you and your students, or other participants, will be more comfortable being videotaped if you and they are accustomed to seeing the camera in the classroom. Again, it becomes part of the daily routine or setting.

Interviews

An alternative to observing people is to directly ask them questions. This can be accomplished in several ways. **Interviews** are conversations between the teacher-researcher and participants in the study in which the teacher poses questions to the participant (Schmuck, 1997). Interviews can be conducted with individuals or with groups. It is best to prepare an **interview guide**, containing either specific or general questions to be asked, prior to conducting any interviews.

Figure 5.1 A sample fieldnote page, showing the left column for actual observations and the right column, used for preliminary interpretations.

Obs. #3 June 10 10:15–11:00	< Observations >	< Observer's Comments (OC) >
Time ↓	There were very few forms of interactions between the children and the teachers. The children were playing; behaving, for the most part. One of the teachers was pushing two girls on swings and the other teacher was sitting near the wading pools, watching the children. Carol said several things to certain children. She repeatedly used phrases such as, "Don't do that," "Don't throw water," "Don't throw that in the pool," and "You're gonna break the sprinkler . . . don't do that!"	I don't think that, in the entire time I was there today, I heard one positive comment or saw one positive gesture. It seemed that the teachers were in only a supervisory role. All they appeared to be doing was supervising the behavior and actions of the children in order to prevent accidents or injuries. I'm not saying that this is wrong; on the contrary, it is necessary when conducting an activity of this nature, especially with very young children. I just expected to hear some positive behaviors being praised in addition to the negative being addressed.
	Several children came close to hurting themselves and/or others. One three-year-old girl tried to pour water over the head of a one-year-old. Two boys were throwing beach balls into the pool and inadvertently hitting smaller children who were playing in the pool.	I began to wonder if this type of activity (i.e., supervisory in nature) did not permit the use of many positive comments. Maybe these teachers leave those types of positive reinforcement for classroom activities. Perhaps activities that require quicker thought and action on the part of the teachers–in order to prevent children from being hurt, or worse–don't allow for positive comments or identification of children to model positive behaviors.
	The children continued to play in the pools, the sprinkler, and the swings. I observed very little verbal interaction between the teachers and the children. Initially, most of what I heard came from Carol. She made several comments to the children, such as "Don't do that" and "You need to ride that bike over there." Carol's daughter picked up a garden hose and began playing with it. Twice Carol told the girl to stop playing with the hose and put it down, but to no avail. The third time she spoke to her, she said, "You better put that down or it will turn into a snake and bite you."	Carol's comment was not in jest. She said it with a firm tone in her voice. I didn't like hearing this. I was always taught never to threaten children, regardless of their age and regardless of how idle the threat. I find myself expecting to see and hear this kind of behavior from Carol and not from Marilyn, as I have not yet heard her say something of this nature.

Similar to observations, interviews are typically classified as being structured, semistructured, or open-ended. In a **structured interview**, the researcher begins with an interview guide consisting of a specific set of predetermined questions. Those questions—and *only* those questions—are asked of each person being interviewed. This is typically done for the sake of consistency. Interestingly, consistency is usually not a concern when collecting

qualitative data; it is typically more desirable for the researcher to have some flexibility and to be able to ask clarifying questions (not initially included on the interview guide), to pursue information not initially planned for, and to seek different information from different people (Leedy & Ormrod, 2005).

When gathering truly qualitative data, interviews are probably best conducted following semistructured or open-ended formats. In **semistructured interviews**, the researcher asks several "base" questions but also has the option of following up a given response with alternative, optional questions that may or may not be used by the researcher, depending on the situation. When developing interview guides, it is best to keep your questions brief, clear, and stated in simple language (Johnson, 2005; Schwalbach, 2003). For example, if we were interviewing students regarding their opinions of our school, we might ask the following questions, where the italicized questions represent the optional, follow-up, probing questions:

- What do you enjoy most about this school?
  Why do you enjoy that aspect so much?
  Do you think other schools have this particular benefit?
- What are your favorite academic subjects?
  Why is that your favorite subject?
  Do you have any others?
  What about extracurricular activities? Are there any that you participate in?
  Which are your favorites? Why?
- What do you like least about this school?
  Why do you like that so little?
  Is there anything that the principal or teachers could do to improve that aspect?

The semistructured interview guide that I used in my positive reinforcement study is shown in Figure 5.2, and a portion of the transcript from one interview I conducted is shown in Figure 5.3.

Open-ended interviews provide the respondent with only a few questions, very broad in their nature. The intent is to gather very different kinds of information from different individuals, depending largely on how each interprets the questions. For example, an open-ended series of interview questions about school climate might include the following:

- What does "school" mean to you?
- What do you like about school?
- What do you dislike?

As mentioned earlier, interviews are not only conducted with individuals but also with groups. A **focus group** is the name given to simultaneous interviews of people making up a relatively small group, usually no more than 10 to 12 people (Leedy & Ormrod, 2005). The length of this type of interview typically lasts between 1 and 2 hours. Focus groups are especially useful when time is limited and because people often are more comfortable talking in a small group, as opposed to individually. Furthermore, interactions among the focus group participants may be extremely informative due to the tendency for people to feed off others' comments. However, when conducting a focus group interview, it is important to ensure that

Figure 5.2 A sample semistructured interview guide.

Semistructured Interview Guide

INTERVIEW WITH THE DIRECTOR

- What type of training and/or certification is held by your classroom teachers?
- Do you have any advice or suggestions for giving positive reinforcement, as discussed with your teachers?
  - How have those suggestions been received by your teachers?
  - Have they attempted to implement them?
- What do you see as acceptable forms of positive reinforcement for children in your school?
- What do you think the meaning of positive reinforcement is for you?
  - Do you think it is the same for your teachers? Why or why not?
  - Do you think it is the same for your students? Why or why not?

INTERVIEWS WITH THE TEACHERS

- Has your director ever provided you with suggestions for giving positive reinforcement?
  - If so, have you used any of them?
  - To what extent have they been successful?
- What do you see as acceptable forms of positive reinforcement for children?
- What do you think the meaning of positive reinforcement is for you?
  - Do you think it is the same for your students? Why or why not?

Figure 5.3 Portion of a transcript from a semistructured interview, using the guide shown in Figure 5.2.

CM: How would you describe positive reinforcement? How would you define that, or what does that mean to you?

"Carol": Positive reinforcement means not yelling at the children. It means talking to them in a positive way. Sometimes you can lose your temper. I try not to use time-out a whole lot. I give them choices. If you're going to throw the blocks, then you're going to pick them up. If you're going to hit someone in the head with that toy, then you're going to go apologize to them. And tell them the difference between right and wrong instead of, . . . take for instance E., who likes to throw toys at everybody. Instead of putting him in the corner and my picking up all the toys he's thrown, I make a game out of it. Instead of "E., pick them up, pick them up," we count them as we put them in. So he's still having to do what he did—you know having to clean up his mess—but we're making a game out of it. Instead of "this was wrong and you're going to sit in the corner for this."

CM: So they don't see it so much as a punishment. Rather, you try to turn it into something constructive?

"Carol": Right. Like this morning, he punched a little girl in the face, and Gail and I both agreed that he needs to sit out of the group for a little while.

CM: So it really depends on the situation? It would be hard to take that situation and turn it into something positive.

"Carol": Right. It depends on what they've done and if they keep doing it all day long. Then they need time away. That's why we have that carpet out there. If the child needs to leave the room and get away from the other children for 5 minutes, they go out and sit on the quiet rug.

Figure 5.4 Sample of guiding questions used for a focus group interview.

1. (a) What were your overall perceptions of the process used to gather student feedback on your teaching?
   (b) What aspects of the process did you like?
   (c) What aspects did you dislike?
2. (a) How was the feedback you received useful to you?
   (b) How was the feedback not useful to you?
3. (a) What changes have you made to any of your teaching behaviors as a result of the student feedback?
   (b) What behaviors, if any, are you considering changing in your teaching as a result of the student feedback?
4. (a) What unanticipated benefits did you experience as a result of this process of collecting student feedback?
   (b) What negative consequences did you experience as a result of this process of collecting student feedback?
5. (a) Is this method, that of using rating scales, the most appropriate way to collect student feedback?
   (b) What method(s) might work better? Why?
6. (a) For what specific school situations or student groups would this method of collecting student feedback not be appropriate?
   (b) What could be changed in order to make it more suitable in this context or to these students?
7. (a) Is this process feasible for teachers to conduct on their own?
   (b) If not, what would need to be changed in order to make it more feasible?
8. (a) How often should this information be collected from students?
9. (a) What specific things could be changed in order to improve this process of collecting student feedback?
10. (a) Based on your experience, will you continue to collect student feedback in this manner?
    (b) If not, will you continue to collect this information but do so by using a different method? Can you describe that method?

* Upon completion of the above questions, explain to the participants that the meeting is about to end. Ask them to take a moment and think about what has been discussed. Then, one by one, ask them if they have any additional comments. If necessary, explore relevant or new comments in greater depth.

each participant is provided with the opportunity to speak and share her or his perspective (Mills, 2003). There can be a tendency for one or two individuals to dominate the discussion; it is the responsibility of the teacher-researcher to closely monitor the discussion in order to prevent this from happening. The set of guiding questions I used for a study incorporating data collected via a focus group is provided in Figure 5.4.

Qualitative data may also be collected via the use of e-mail interviews (Mills, 2003). With schools becoming increasingly networked, teacher-researchers can easily collect data from colleagues, parents, and students by sending out a series of questions in an e-mail message. One benefit of doing so is that when the respondent replies to your e-mail questions, the transcription of the interview has already been done for you. However, you must be cautious of possible ethical complications and realize that e-mail responses are not necessarily anonymous or

confidential (Mills, 2003). Other individuals who may have access to a server may be able to intercept e-mail responses from targeted respondents.

Hubbard and Power (2003) also remind teacher-researchers not to forget about the value of **informal interviews**, that is, those that are spontaneous, that take place throughout the data collection process, and that are typically part of the daily interactions with students in a classroom setting. Teachers are constantly asking students questions, trying to gather various types of information from them.

Schmuck (1997) provides a discussion of the relative advantages and limitations of conducting interviews as part of action research studies. Advantages include the fact that interviews permit the teacher-researcher to probe further and ask for clarification in a participant's response to a given question. In addition, data can be collected—and, therefore, preserved—through the use of audio- and videotapes, although you want to be sure that individuals being interviewed are not made to feel uncomfortable by the presence of an audio or video recorder. Finally, for respondents who cannot or who are unwilling to share their thoughts, feelings, or perceptions in writing, sitting down and carrying on a conversation about them is often a reasonable alternative. On the other hand, interviews can be extremely time consuming. Not only does it take time to collect data from individuals during a verbal conversation, but before the data can be analyzed, the interviews must be transcribed so that the responses can be read and processed. The general rule of thumb that I learned in my graduate school days is that for every hour of audiotaped interview, you can expect approximately 8-9 hours of transcription work, depending on the quality of the recording. Other limitations of interviews include the fact that respondents are not able to retain their anonymity. Many people are simply uncomfortable with a tape recorder lying on the table between them and the interviewer. Finally, respondents often fear that something they have said may be used against them at some point in the future. An additional responsibility of the teacher-researcher is to put the mind of the interviewee at ease about such possibilities.

Journals

Data journals may be kept by both teachers and students and provide valuable information into the workings of a classroom (Mills, 2003). In a way, **student journals** provide information similar to homework to the teacher, in that teachers can gain a sense of students' daily thoughts, perceptions, and experiences in the classroom. **Teacher journals** can similarly provide teacher-researchers with the opportunity to maintain narrative accounts of their professional reflections on practice. They truly become an ongoing attempt by teachers "to systematically reflect on their practice by constructing a narrative that honors the unique and powerful voice of the teachers' language" (Mills, 2003, p. 68) by reflecting not only observations but also the feelings and interpretations associated with those observations.

Class journals are another means of incorporating journaling into your action research data collection. A **class journal** is a less formal version of a student journal. Johnson (2005) suggests that a blank notebook be passed around the class on a periodic basis or put in a learning center for an extended amount of time. Students are encouraged to enter their thoughts, ideas, perceptions, feedback, or other forms of response, such as pictures or diagrams, as they wish. Teachers may want to provide some sort of guidelines for making entries into the class journal so that it does not become a "quasi-teacher-approved" form of graffiti that may be offensive to other students (Johnson, 2005).

Existing Documents and Records

Often, action research necessitates the gathering of data that already exist. These data are essentially anything collected for a reason *other* than the action research study but are now being used as data for the study. These **existing documents and records** might take several forms, including (at the individual student level) curriculum materials, textbooks, instructional manipulatives, attendance records, test scores, previous grades, discipline records, cumulative folders, and (at the school or district level) attendance rates, retention rates, graduation rates, newspaper stories about school events, minutes from faculty or school board meetings, and standardized test scores perhaps disaggregated by grade level, gender, or ethnicity (Johnson, 2005; Mills, 2003; Schmuck, 1997). However, a word of caution is in order: Whenever using existing data, it is critical to make sure to follow your school district's approved procedures for securing access to these various types of data and that you use and report the results of any analyses in an ethical manner (Johnson, 2005).

Collecting existing data on students can sometimes become overwhelming—there may be *so much* information you want to collect. The dilemma often faced by teacher-researchers is how to organize that information. A nice organizational tool that can aid in your organizational efforts is to record data on a common data form. In other words, you develop for your specific purposes and use a single form as a means of compiling various types of information, as opposed to having a conglomeration of loose papers stuffed in a file folder, for example. Several years ago, I was involved in a study that examined student attendance records, reasons for school absences, discipline referrals, and referrals for special programs and social services. That was a good deal of information to collect on each student. However, I designed a data collection form (see Figure 5.5) that allowed us to collect this wide variety of information on a single form, thus organizing and encapsulating it at the same time.

One possible type of existing "data" that should not be overlooked is that of classroom artifacts. **Classroom artifacts** include any written or visual sources of data, contained within the classroom, that contribute to our understanding of what is occurring in classrooms and schools (Mills, 2003). This source of existing data primarily incorporates work done by students as part of their schoolwork but is now used as data for action research purposes (Hubbard & Power, 2003). Items such as student portfolios or products resulting from the administration of a performance-based assessment, or less formal artifacts such as students' responses to a teacher's request to explain the solution to a constructed-response mathematics problem, can serve as prime examples of student-produced classroom artifacts.

Characteristics of Qualitative Data: Accuracy, Credibility, and Dependability

When collecting data for action research studies, it is important for teacher-researchers to ensure the quality of their data. If data collected for the study are imprecise, or if the researcher has actually measured something other than what was intended to be measured, at a minimum the data will be inaccurate and misleading. The larger concern here is that if the action research study is continued to its logical "end," the results of the study will follow suit: they too will be inaccurate and misleading. If that occurs, you have essentially wasted your time—not to mention that of your colleagues, students, parents, and anyone else involved in your study.

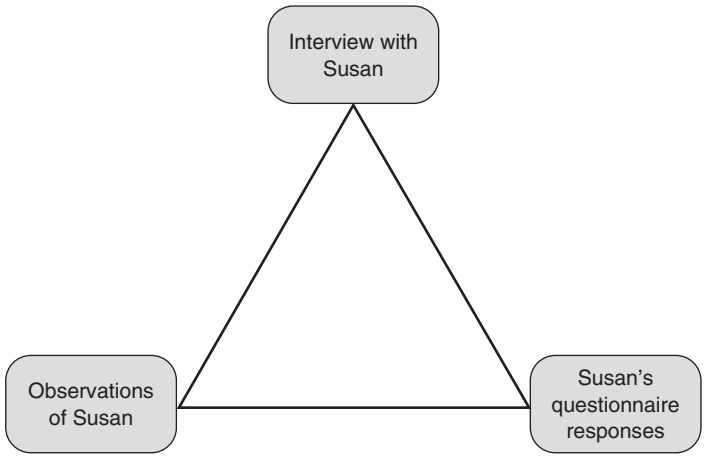
Figure 5.5 Sample of a data collection form for existing student data.

Student Name	*Days Absent*	*Reason for Absences*	*# Disc. Referrals*	*Reasons for Disc. Referrals*	*Referral for Special Program? (Y/N)*	*Referral for Social Services? (Y/N)*	*Retained? (Y/N)*

Validity of research data deals with the extent to which the data that have been collected accurately measure what they purport to measure (i.e., that which we intended to measure) (Mills, 2003). When dealing with the validity of qualitative data, researchers are essentially concerned with the **trustworthiness**—e.g., the accuracy and believability—of the data. Trustworthiness is established by examining the credibility and dependability of qualitative data. **Credibility** involves establishing that the results of qualitative research are credible or believable from the perspective of the participant in the research (Trochim, 2002c). On the other hand, the concept of **dependability** emphasizes the need for the researcher to account for the ever-changing context within which research occurs. The researcher is responsible for describing the changes that occur in the setting and how these changes affected the way the researcher approached the study (Trochim, 2002c).

There are three common practices, typical aspects of any qualitative research study, that can help to ensure the trustworthiness of your data. The first of these is triangulation, or the use of

Figure 5.6 Triangulation of three sources of data.



multiple data sources, multiple data-collection methods, and perhaps even multiple teacher-researchers in order to support the ultimate findings from the study (Glesne, 1999; Hubbard & Power, 2003). A given finding is supported by showing that independent measures of it tend to agree with each other or at least do not directly contradict each other (Hubbard & Power, 2003). For example, when you observe Susan *actually doing* something that she has *told* you in an interview that she does and that is also indicated on an open-ended questionnaire (see Figure 5.6), you likely will have more confidence in concluding that it is probably an accurate depiction of Susan's practice. In other words, your interview data has been supported by your observation data and by the questionnaire responses. Had any of the three sources of data contradicted each other, you likely would have arrived at a different conclusion, perhaps that Susan was telling you what you wanted to hear, although in reality she did not practice it.

A second practice that can help to ensure the quality of your data is known as **member checking**. This procedure involves the sharing of interview transcripts, analytical thoughts (such as observation notes with observer's comments), and drafts with the participants of the study. The purpose of sharing these data sources is to make sure that you have represented your participants and their ideas accurately (Glesne, 1999). A third and final procedure involves **prolonged engagement and persistent observation**. The idea here is that the more time you spend "in the field," so to speak, the more you are able to develop trust with and get to know your participants, learn the culture of their setting (whether it be a classroom or school building), and observe patterns of behavior to the point of being routine (Glesne, 1999). Observing and/or interviewing only once or twice will not afford you this luxury.

QUANTITATIVE DATA COLLECTION TECHNIQUES

In contrast to qualitative data, quantitative data are *numerical.* Anything that can be quantified (i.e., counted, calculated, tallied, rated, etc.) can be considered quantitative data. This includes not only items that can be counted but also ratings of one's feelings, attitudes,

interests, or perceptions on some sort of numerical scale. Quantitative data collection techniques include surveys, questionnaires, checklists, and rating scales, as well as tests and other more formal types of measurement instruments. Generally speaking, quantitative data collection techniques are more efficient, in that you can collect data from numerous individuals simultaneously. However, the depth of that data does not begin to compare to that resulting from the use of qualitative techniques.

Surveys, Questionnaires, and Rating Scales

The term "survey" refers to a collective group of quantitative data collection techniques that involve the administration of a set of questions or statements to a sample of people. Surveys may be administered orally—which then makes it a type of interview, although the resulting data are numerical instead of narrative—or in written form. Surveys that are administered in written form, where the researcher asks participants to answer a series of questions or respond to a series of statements and then return it to the researcher, are known specifically as questionnaires. Surveys and questionnaires permit the teacher-researcher to gather a lot of—as well as a variety of—information relatively quickly (Johnson, 2005). There is not much that limits your use of surveys and questionnaires. They may be simple or complex; they can be comprised of **open-ended questions** (where individuals provide their own responses) or of *closed-response* rating scales (where individuals simply select their response from a set of options provided to them), or may even contain a combination of the two types of questions or statements. Analysis of responses to open-ended items—which are, admittedly, more qualitative in design than are closed-response items—may often reveal unexpected thoughts and feelings from students, the likes of which you may not have been able to anticipate in order to develop closed-response items to address those particular thoughts and feelings (Schmuck, 1997).

A closed-response question or statement provides the respondent with a number of choices from which to select. For example, consider the following item:

> What is your favorite subject in school?

1. English
2. Mathematics
3. Science
4. Social Studies

Students would be instructed to select one of the four possible responses. This type of question is easily quantifiable; you simply count the number of students who select each option. Furthermore, it is relatively easy to report the "results" of this item. You might summarize your data and conclude the following:

35% of students responding prefer science

25% of students responding prefer mathematics

25% of students responding prefer English

15% of students responding prefer social studies

It is important to realize that this type of question may be misleading or controlling (Johnson, 2005). If, in our example, the favorite subject of a given respondent is a foreign language class, how is that person supposed to respond to the question? Any option that person selects will actually provide inaccurate information. One alternative is to anticipate such an occurrence by revising the item to read as follows:

> What is your favorite subject in school?

1. English
2. Mathematics
3. Science
4. Social Studies
5. Other: _____

Open-ended items allow the respondents to provide a seemingly limitless number of responses. For example, we could have reworded our "favorite subject" question as an open-ended question by simply asking

> What is your favorite subject in school?

Here we might get a wide variety of responses. It is then the responsibility of the researcher to "analyze" the resulting data by grouping similar items together and then tallying the number of responses in each category. The result might look like this:

26% of students responding prefer science

25% of students responding prefer English

15% of students prefer geometry

15% of students responding prefer social studies

10% of students responding prefer mathematics

5% of students prefer art

2% or students prefer physical education or health

2% responded, "I don't know" or "I don't have a favorite subject"

Obviously, this form of the question provides a more accurate sense of what students really like. The only problem associated with asking open-ended items like this is that you have the sometimes messy task of grouping responses into similar categories before you can count the responses (Johnson, 2005).

The main difference between a survey, or questionnaire, and a rating scale is that surveys are more appropriate for content-based types of questions (similar to our example above), whereas rating scales are appropriate when asking individuals to respond to a set of questions where their response indicates the strength (e.g., the extent of agreement, level of frequency, degree of understanding) of that response (Johnson, 2005). Rating scales can be used very effectively to measure students' attitudes, perceptions, or behaviors. There are two main types of scales that appear in items on a rating scale: Likert and Likert-type scales. A **Likert** (pronounced "lick-kert") **scale** begins with a statement and then asks individuals to respond on an agree-disagree continuum. The Likert scale typically ranges from strongly agree to strongly disagree. I typically recommend using a five-point scale, with the five points defined as follows:

1 = strongly disagree

2 = disagree

3 = no opinion

4 = agree

5 = strongly agree

There tends to be quite a bit of disagreement among those with expertise in conducting research through the use of surveys regarding the appropriateness of including a neutral point on a scale. By including it, you allow your respondents to indicate that they truly are neutral or have no opinion, if in fact that is the case for them. However, if provided with a neutral option, there is a tendency for people *not* to think much about how they truly feel; they simply select the neutral option, which may not represent their true belief (i.e., the data they provide is inaccurate). On the other hand, if individuals *truly* are indifferent or have no opinion, and you do not provide this option—because you are operating under the assumption that no one is truly neutral about anything—you "force" them to choose something that they do not really believe, thus providing inaccurate data once again. There is no right or wrong when it comes to deciding on the inclusion of a neutral point on your rating scale. However, you should consider the implications of both including and excluding such a point and then design your scale accordingly. Figure 5.7 presents a portion of a rating scale that I used in a study that focused on students providing their teachers with feedback on their classroom teaching. Notice the format of the Likert-scaled items. Also notice that a higher number corresponds to a higher level of agreement with a given statement.

A similar type of scale is a **Likert-type scale**. This type of scale also exists on a continuum, but something other than extent of agreement is being measured. For example, a Likert-type item might require participants to respond on a scale that examines quality ("excellent . . . poor"), frequency of occurrence ("always . . . never"), or level of comfort ("very comfortable . . . not at all comfortable") (Mertler & Charles, 2005). An example of a Likert-type scale, used in a study of prekindergarten-to-kindergarten transitions, is shown in Figure 5.8.

Figure 5.7 Portion of a rating scale instrument, depicting a Likert scale.

STUDENT EVALUATION OF TEACHERS AND TEACHING TECHNIQUES (SE3T)

The purpose of this questionnaire is for you to help your teachers to improve. Several statements about your teacher are listed below. Please circle the number, using the code below, that describes how much you agree with each statement. Your responses will be anonymous; please do not place your name anywhere on this form. Please respond to each statement as honestly as you possibly can and by circling only one number for each statement.

1	2	3	4	5
Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree

1.	My teacher tells me when I do good work.	1	2	3	4	5
2.	My teacher encourages me to ask questions when I don't understand what's going on in class.	1	2	3	4	5
3.	My teacher tells us why the things we are learning are important.	1	2	3	4	5
4.	My teacher makes it clear how grades are determined.	1	2	3	4	5
5.	I really pay attention in this class.	1	2	3	4	5
6.	My teacher is fair when students misbehave.	1	2	3	4	5
7.	My teacher teaches things in an order that makes sense.	1	2	3	4	5
8.	My teacher takes time in class to help students.	1	2	3	4	5
9.	This class is challenging to me.	1	2	3	4	5
10.	My teacher gives fair tests.	1	2	3	4	5
11.	It is important to me to learn the material in this class.	1	2	3	4	5
12.	My teacher knows how to handle students who disrupt class.	1	2	3	4	5
13.	My teacher uses my ideas in class.	1	2	3	4	5
14.	My teacher explains assignments well.	1	2	3	4	5
15.	My teacher returns tests without much delay.	1	2	3	4	5
16.	My teacher has a sense of humor.	1	2	3	4	5
17.	I am sometimes confused about what's going on in this class.	1	2	3	4	5
18.	My teacher encourages good student behavior in class.	1	2	3	4	5

I want to mention one more thing about using surveys and rating scales with students. Teacher-researchers need to be sure that the various aspects—not just the reading level—of the instrument are appropriate for the age or grade level of students. Although I recommended earlier that a five-point scale is typically appropriate, one could see how that might create difficulties for young children—they obviously would not be able to discriminate between adjacent points on the scale. However, do not shy away from using such data collection instruments with younger children. You would likely provide fewer options on the scale and perhaps even use graphics for the children to respond to. Several years ago, I was part of a research team that attempted to "survey" kindergarten students as part of the prekindergarten-to-kindergarten

Figure 5.8 Portion of a rating scale instrument, depicting a Likert-type scale.

Teacher Assessment of Student Adjustment to School

Directions: Please list all students and rate each student on the eight characteristics listed as they relate to the beginning of school. Use the numbered scale listed below. In addition, feel free to add any comments that would aid in describing the adjustment of the students.

1 2 3 4 5
|----------|----------|----------|----------|----------|
Not at All — Some of the Time — All of the Time

Student Name	**Adjustment Indicators**							
	Fearful	Relates Well to Peers	Complains of Illness	Cries Easily	Exhibits Self-Confidence	Frustrated	Intimidated	Appears Happy

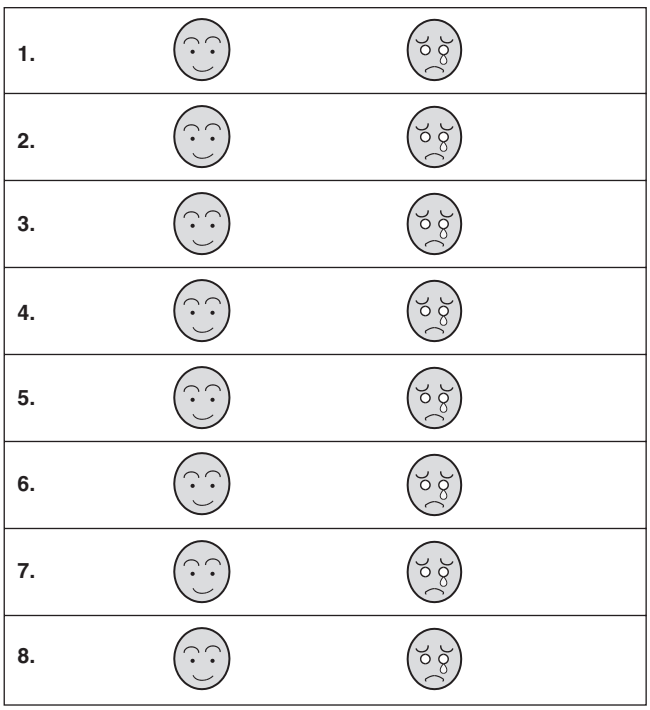
transitions study. We had the teachers read the statements to the children, then asked them to put an *X* through the face that represented how they felt (see Figure 5.9).

Figure 5.9 Triangulation of three sources of data.

Student Self-Assessment Survey

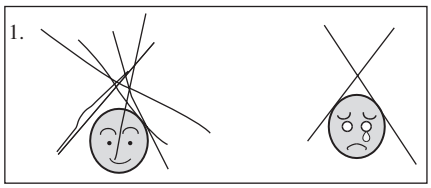
Directions: Ask your students to think back to the very first day of Kindergarten. Read each item to your students and instruct them to respond at the appropriate space on the Student Self-Assessment Survey Response Form.

1. How did you feel about the first day of Kindergarten?
2. How did you feel about leaving home on the first day of Kindergarten?
3. How did you feel about meeting a new teacher this year?
4. How did you feel about making friends with your new classmates?
5. How did you feel about playing on the playground?
6. How did you feel about eating lunch in the cafeteria?
7. How did you feel about the activities you did on the first day of school?
8. How did you feel about leaving school at the end of the first day?



Unfortunately, the children had no idea—and our explanations did not help at all—what the numbers were for. They were instructed to locate the number 1 on their response sheet, as the teacher read statement number 1, and then place their *X* on the appropriate face. After the

first few statements, we realized that they were simply placing the ***X*** over the same faces in the first row. Several of the children had response sheets that looked like this:



Obviously, you can see the problems that this created with respect to the accuracy of our data! On the spur of the moment, we decided to revise the nature of the response sheet and came up with what you see in Figure 5.10. Using this format, we could direct their attention to the box with a certain image in it and have them place the response only in that box.

The advantages of surveys and rating scales include the fact that they are very effective at gathering data concerning students' attitudes, perceptions, or opinions. They are essentially written versions of structured interview guides, where individuals respond to a specific set of questions in writing, as opposed to responding orally. Rating scales and other closed-response items can be answered, and the responses can be tallied or counted quickly. Integrating the use of computer software can make this process of tallying even quicker.

There are, of course, also limitations to the use of surveys for action research projects. Analyzing responses to open-ended items can sometimes be time consuming, due to the fact that responses may be ambiguous (Schmuck, 1997). This limitation can be overcome by replacing open-ended items with rating scales or other closed-response items. Another limitation is that if the teacher-researcher is not clear about an individual response, there is no opportunity or mechanism for asking respondents to clarify their answer, as with interviews.

At this point, I would like to offer several suggestions—adapted from several sources (Johnson, 2005; Mills, 2003; Schmuck, 1997; Schwalbach, 2003)—regarding the development and use of surveys and rating scales as means of collecting action research data. When developing a new instrument, it is important to do the following:

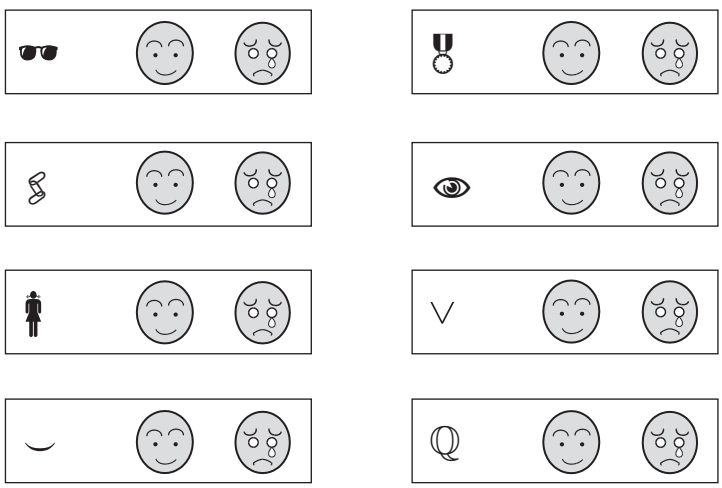
- Each item should focus on a single idea or concept.
- Do not use too many questions or questions that are not necessary or are repetitive.
- Keep the length of the survey brief and the reading level relatively easy; failing to do so often results in respondents not completing the instrument or providing you with inaccurate information.
- If you are designing a rating scale, you should keep the response scale consistent throughout the survey. Otherwise, respondents can become confused or may provide you with inaccurate data.
- Consider using both closed-response and open-ended items, in order to realize the benefits of both.
- Do not use *leading* questions; a good survey or rating scale is one that contains objective items.
- Always proofread your survey—and perhaps have someone else proofread it too—before you administer it to your participants.

Figure 5.10 Revised version of instrument presented in Figure 5.9.

Student Self-Assessment Survey

Directions: Ask your students to think back to the very first day of Kindergarten. Read each item to your students and instruct them to respond at the appropriate space on the Student Self-Assessment Survey Response Form.

1. How did you feel about the first day of Kindergarten?
2. How did you feel about leaving home on the first day of Kindergarten?
3. How did you feel about meeting a new teacher this year?
4. How did you feel about making friends with your new classmates?
5. How did you feel about playing on the playground?
6. How did you feel about eating lunch in the cafeteria?
7. How did you feel about the activities you did on the first day of school?
8. How did you feel about leaving school at the end of the first day?



Checklists

A checklist is a list of behaviors, characteristics, skills, or other entities that a researcher is interested in investigating (Johnson, 2005; Leedy & Ormrod, 2005). The primary difference between a checklist and a survey or rating scale is that checklists present only a dichotomous set of response options, as opposed to some sort of continuum. Instead of indicating the extent, degree, or amount of something, checklists enable the teacher-researcher to indicate simply *if* the behavior or characteristic is observed or present or if it is not observed or present. Checklists are quicker for the teacher-researcher to use than are surveys and rating

Figure 5.11 A sample student checklist, looking at independent reading at the elementary level.

Independent Reading Checklist

Student Name: _____

Grade: _____

Date: _____

Independent Reading Trait	***Observed?***	***Not Observed?***
Respect others	☐	☐
Stays on task	☐	☐
Chooses appropriate books	☐	☐
Stays focused on the story	☐	☐
Understands various elements of the story	☐	☐
Thinks about the story	☐	☐
Is able to answer questions about the story	☐	☐
Thinks about the characters	☐	☐
Is able to answer questions about the characters	☐	☐

scales; however, they provide data that are not nearly as detailed as those resulting from the use of rating scales.

If you are observing students, of any age, and are using a checklist to record behaviors, you will want to keep the list of behaviors or characteristics to a manageable number. Otherwise, you can become overwhelmed with the sheer volume of things you must observe and record on the checklist. A sample student checklist is presented in Figure 5.11.

Tests and Other Formal Instruments

Tests—whether they be standardized tests or teacher-developed, classroom tests—can also be used as sources of quantitative data. It is important to realize that these "formal" data collection instruments would also be considered "existing records," since they are administered to students as a regular part of classroom instruction and district-level accountability. Other formal assessment instruments that might be included here are scores on homework assignments and quizzes, and final semester or course grades. If they are included as data in your action research study, it is important that they not be the only source of data (Johnson, 2005).

In addition, there may be instances—depending on the nature of the questions and design of your action research study—where you might design a pretest and posttest specifically to

measure changes that (hopefully) occur as a result of exposing students to some sort of instructional "treatment" or intervention. In this case, these may not be considered as existing data, since you designed and administered them with a specific research-related purpose in mind. They do, however, remain an important source of quantitative data.

Characteristics of Quantitative Data: Validity and Reliability

Validity is an essential quality in quantitative research data and has to do with whether the data are, in fact, what they are believed or purported to be—in other words, did we *actually* measure what we intended to measure, based on the focus of our research? Though any data that you might collect may be entirely accurate, the critical factor is whether or not it is appropriate and accurate *for your purposes* (i.e., does it enable you to accurately answer your research questions?). For example, imagine that a reading teacher uses the results from the reading portion of a standardized test to group her students into above-average, average, and below-average reading groups. Then imagine that a social studies teacher uses those same reading scores to identify students who he believes would be successful in an advanced placement history course. The first interpretation and use of the scores is valid; the second is not. In terms of the social studies teacher's use of the data, it was invalid for the purpose for which it was used. The determination of the validity of data ultimately has a substantial effect on the interpretation of that data, once they have been analyzed, and the subsequent conclusions drawn from those results (Mertler & Charles, 2005).

Presently, validity is seen as a unitary concept (AERA, APA, & NCME, 1999), combining that which has been previously described as four distinct types of validity: content, concurrent, predictive, and construct. It is defined as the "degree to which all the accumulated evidence supports the intended interpretation of test scores for the proposed purpose" (p. 11). Validity of quantitative data can be determined through the examination of various sources of evidence of validity. Although similar to the four outdated types of validity, the five sources of validity evidence are unique in their own right (Mertler & Charles, 2005). These five sources of evidence of validity are based on the following: test (or instrument) content, response processes, internal structure, relations to other variables, and consequences of testing. Many of these sources of validity evidence are more appropriate for large-scale testing programs, especially where it is important for the results to be generalizable to much larger populations than simply those individuals included in a research study. Since this is not a purpose or goal of classroom-based action research, I am suggesting that teacher-researchers be most concerned with evidence of validity *based on instrument content.* This source of evidence is based on the relationship between the content addressed on a test, or other instrument used for data collection, and the underlying **construct** (or characteristic) it is trying to measure. For example, assume we wanted to survey students to determine their attitudes toward learning mathematics. We would want to ensure that the questions we asked on the survey dealt directly with various aspects of learning math, not learning in any other subject areas, or questions that were completely extraneous to the construct of "learning mathematics." As another example, consider a test you might administer to students on their understanding of the process of photosynthesis. If you wanted to be able to draw conclusions *specifically* about their understanding of this scientific process, you would need to be sure to ask only questions related to the process. If unrelated questions were also asked of students on the test—and

provided that they contributed to the overall score on the test—interpreting the scores as an indication of their understanding *only* of photosynthesis would not be a valid, legitimate use of those scores. This type of evidence is typically based on subjective, logical analysis of content coverage on the test and can be established by critical review by teachers, as well as by the judgments of experts in the particular content field. In other words, although it is a subjective process, it is important for teacher-researchers to critically examine the individual items and overall content coverage on a survey, rating scale, checklist, test, or quiz in order to ensure that they are measuring what they intended to measure.

Reliability, a second essential characteristic of quantitative data, refers to the consistency of collected data. If you to hear three accounts of a minor car accident from three different individuals, but each account differs as to what happened, who was involved, and what the results were, you would likely have little confidence in any of the versions you have heard. In other words, the accounts (the data) are inconsistent and, therefore, unreliable. If, however, each account is essentially similar, the information you have received is consistent and may be considered reliable. Similarly, if you administer a certain test repeatedly under identical circumstances but find that you get different results each time, you would conclude that the test is unreliable. If, however, you get similar results each time you administer the test, you would consider the results reliable and, therefore, potentially useful for your purposes (Mertler & Charles, 2005).

As with the determination of the validity of quantitative data, there are several methods of determining the reliability of data (Mertler & Charles, 2005), not all of which are appropriate for teachers conducting classroom-based research. Reliability of quantitative data is usually established by correlating the results with themselves or with other quantitative measures. Three different methods are used—test-retest, equivalent forms, and internal consistency. **Internal consistency** is a statistical estimate of the reliability of a test that is administered only once. For this reason, this type of reliability estimate is most useful for classroom teachers conducting research. One of the easiest internal consistency formulas to use is the **Kuder-Richardson formula 21** (also known as KR-21). The resulting statistic will range from 0.00 to 1.00; the closer the value is to 1.00, the more reliable your data are. The formula for calculating KR-21 internal consistency is

$$r = \frac{(K)(SD)^2 - M(K - M)}{(SD)^2(K - 1)}$$

where r is the reliability index, K is the number of items on the test or instrument,—is the mean or average score, and SD is the standard deviation of the scores. Imagine that a test consists of 40 items, that the mean is equal to 27.3, and that the standard deviation is 4.64. The internal consistency reliability for the exam, using the KR-21 formula, is shown below.

$$r = \frac{(40)(4.64)^2 - (27.3)(40 - 27.3)}{(4.64)^2(40 - 1)} = \frac{(40)(21.53) - (27.3)(12.7)}{(21.53)(39)} = \frac{514.49}{839.67} = .61$$

We often think of validity and reliability as two distinct concepts, but in fact they share an important relationship (Mertler & Charles, 2005). It is possible for scores obtained from an instrument to be reliable (consistent) but not valid (measuring something other than what was intended). In contrast, scores cannot be both valid and unreliable—if scores measure what was intended to be measured, it is implied that they will do so consistently. Therefore, reliability is a necessary, but not sufficient, condition for validity. When establishing the validity and reliability of your research data, always remember the following adage: *A valid test is always reliable, but a reliable test is not necessarily valid* (Mertler & Charles, 2005).

WRITING UP ACTION RESEARCH: DATA COLLECTION METHODS

I teach eighth grade, comprehensive science (an integrated life, earth/space, and physical science program) at a middle school in a rural North Florida county. I am a member of a four-teacher team, along with one math, one history, and one language arts teacher. We instruct 130 students who make up the academic team (Team E). Although I teach five science classes a day, I targeted my seventh period class for my research.

This class is made up of thirty-one average and above-average science students. I chose this last class of the day for purely logistical reasons. With only one computer in my classroom, I needed to borrow eleven computers daily from neighboring teachers. Seventh period was the most agreeable period to the other teachers. An extra advantage of using the last period of the day was that students could return the computers after the final dismissal bell and not take valuable class time for this task.

My data was generated by comparing these students' attitudes toward learning science at the beginning of the school year, during my study, and at the conclusion of the study period. The students' attitudes and reactions were documented by the students themselves, by their parents, and by my own observations. Collecting data from three sources allowed for triangulation of the findings in this study. Data triangulation helped reduce the likelihood of error in the findings when similar results are reported from two or more of the sources. I surveyed all of the class members and their parents at the beginning and the end of my study.

Data collected via surveys from three sources; allowed for triangulation

During the first six weeks of school, I reviewed the scientific method, the metric system, scientific measurement, and laboratory safety. At this point multimedia technology was not part of the curriculum. Some hands-on activities were used at this time. The students worked both individually and in groups. To determine each student's level of enthusiasm for learning science, during this time I administered a survey which contained the following questions: How do you like learning science? How have you liked learning science so far this year? How enthusiastic are you about exploring science at home? Students were asked to rate their answers to each question using a scale of 1 to 5. The scale was represented by (1) a very unenthusiastic response, (2) an unenthusiastic response, (3) indifference, (4) an enthusiastic response, and (5) a very enthusiastic response.

Additionally, I sent home parent surveys with each student in order to solicit and record the parents' opinions concerning their child's enthusiasm for learning science. The survey included

two questions: How enthusiastic is your child about learning science? How enthusiastically does your child do science activities at home? I used the same rating scale for the parents that I used with the students.

Description of parent survey

At the beginning of the second six weeks I introduced a unit on oceanography. Oceanography was used as the unit of study primarily because of the number of resource materials available to the students through the media center. It was during this unit that I began to integrate technology into my curriculum. As the unit was introduced I asked my students to look through the oceanography chapters in their textbooks and make a prioritized list of the eleven subtopics in physical and biological oceanography they would like to study. Students were grouped according to their interest as much as possible and were assigned to work in groups of two or three to develop a multimedia presentation that would be used as an instructional tool for the other students.

During this period I began to introduce them to the multimedia computer program, HyperStudio (Wagner, 1994). HyperStudio is a program that allows the user to combine sound, graphics, and animation with text to make creative and entertaining presentations. The introduction of HyperStudio and the development of the student presentations took six weeks to complete.

Throughout the study I observed and made notes as to how the students were working and their reactions to class. These observations were guided by several questions: What problems are the students encountering as they work on their multimedia presentations? Are the students having problems with content? Are there problems working in groups? Are they having problems using the multimedia software? These observations and notes were useful in making sense of any fluctuations I found in the end-of-study student surveys. I was able to discern the source of problems so that content difficulties or friction within groups was not confused with a loss of enthusiasm for technology.

Observations also conducted

At the end of the oceanography unit I had each group of students share their presentations with the rest of the class. After the presentations, each group was asked to comment to the class on how they enjoyed developing their works. I noted these student comments as they were presented to the class. Each student was also asked to make written, individual comments to me, responding to the following questions: What problems did you encounter while you were developing your presentation? What did you learn about your topic while you were developing your presentations? Did you learn from the other students' presentations? Would you like to do another presentation on some other topic in science? Again I surveyed the parents of these students to gain information about their child's interest in learning science. I asked the following questions: Is your child talking about science at home? Is your child eager to share what we are doing and learning in science class? Do you feel that your child is learning science? Why or why not? How enthusiastic is your child about learning science? How enthusiastically is your child doing science activities at home? I again surveyed the students asking the same questions that I had asked in the beginning survey.

Another source of individual student data

From "Effect of Technology on Enthusiasm for Learning Science" by J. Hollis (1995), in S. Spiegel, A. Collins, & J. Lappert (Eds.), *Action Research: Perspectives From Teachers' Classrooms. Science FEAT (Science For Early Adolescence Teachers)*. Tallahassee, FL: SouthEastern Regional Vision for Education. Retrieved October 20, 2004, from http://www.enc.org/professional/learn/research/journal/science/document.shtm?input=ENC-0024322432_ch1.

Two weeks prior to my starting date, a video camera was placed in my first period classroom and left on so that the students would become comfortable in the presence of the camera in the room. Students were given numbers on construction paper and asked to hold on to them for later use.

Data collected through the use of video recorder and brief surveys administered to students

On day one the first period class was videotaped for the first time. At the close of the period students were asked to complete a four-question survey. They were asked not to use their names, but instead, they were asked to use a number that was given to them earlier. I jotted down notes on how the class session went in a teacher journal.

More videotaping and student surveys

The week continued with the second taping three days later. Student surveys were filled out for the entire week. Entries were made in the teacher journal whenever I could remember. This turned out to be about three times during that first week.

During the second week the class was taped on Monday and Thursday. At the end of the second week modifications to the student survey were made on questions I and 3 due to mixed responses given by students. The modified student survey questions were:

Questions appearing on the final version of the student survey

1 Did you share something in class today? Yes/No

2. If yes, did you share with:
   a) students only;
   b) the teacher;
   c) a group of students; or
   d) a group of students and the teacher.

3. Did you ask a question today? Yes/No

4. If yes, did you ask a question of:
   a) a student only;
   b) teacher only;
   c) group of students only; or
   d) group of students and the teacher.

I continued to tape my first period science class twice a week for a total of five weeks. Student surveys were given to all students on a random basis throughout the five-week period. Journal entries were made daily.

From "What Patterns of Teacher-Student Communication Exist in My Classroom?" by E. Graham (1995), in S. Spiegel, A. Collins, & J. Lappert (Eds.), Action Research: Perspectives from Teachers' Classrooms. Science FEAT (Science For Early Adolescence Teachers). Tallahassee, FL: SouthEastern Regional Vision for Education. Retrieved October 20, 2004, from http://www.enc.org/professional/learn/research/journal/science/document.shtm?input=ENC-002432-2432_ch4. Reprinted with permission.

Reprinted with permission.

ACTION RESEARCH PORTRAIT 1

Enhancing Academic Performance Through Improved Classroom Assessment

Recall that the purpose of this action research study is to improve teachers' classroom-based assessments in an effort to improve student achievement.

The teachers who make up Team North at Jones Middle School felt very comfortable about the way in which their data would fit into the pretest-posttest control group design that they selected for their action research study. They developed consent forms that both the students in Team North and Team East, as well as their parents, signed. The forms requested permission for the students' fall (October) and spring (March) test scores, resulting from the two administrations of the statewide proficiency test, to be used for an additional purpose—their action research study.

Approximately 4 weeks after each administration of the test, the individual student test reports came back to the school. For each of the students in the two teams, the four Team North teachers pulled the test report from the student's cumulative folder in the main office. From the test report, they recorded the scaled score (ranging from 200 to 500) for each of the four main subtests: language arts, mathematics, science, and social studies. They recorded the scores for each subtest (where "1" indicated the fall test scores and "2" indicated the spring scores), along with each student's identification number and team membership (where Team North was coded "1" and Team East "2") in a spreadsheet, which looked like this:

student_id	***group***	***la_1***	***math_1***	***sci_1***	***ss_1***	***la_2***	***math_2***	***sci_2***	***ss_2***
00013579	1	355	410	400	450	370	480	410	460
00024680	1	350	250	420	380	370	260	430	420
00012345	2	410	450	460	390	460	440	460	400
00098765	2	380	290	400	430	410	310	380	280

They double-checked each of the students' test scores across the two administrations of the proficiency test for accuracy of entry into the database. When all of the scores had been verified, they prepared their data (and themselves) for the next step—data analysis!

ACTION RESEARCH PORTRAIT 2

Improving Reading Comprehension in a Title I Program

Recall that the purpose of this action research study is to improve students' reading comprehension skills within a Title I context.

In order to address her initial research question—which proposed to examine differences in students' reading comprehension skills following the use of revised teacher-developed comprehension items, based on pretest and posttest diagnostic test scores—Kathleen needed to select an appropriate and valid measure of reading comprehension. After reviewing the various diagnostic tests with which she was familiar and had

experience administering, Kathleen selected the Woodcock Reading Mastery Test–Revised (Form H) to administer to her reading students in September and in May. From the resulting student score reports, she would extract the Reading Comprehension Cluster score, which appears as a percentile rank. An average score on this subtest is the 50th percentile; Kathleen's upper-elementary students typically score near the 35th percentile. She obviously hoped to improve that performance over the course of the school year.

Kathleen's second research question dealt with the perceptions held by both her students and herself regarding the students' reading comprehension skills. She proposed to collect two forms of data to enable her to address the nature of those perceptions. First, she would conduct daily observations of her students and record both what she saw and any analytical thoughts she may have while conducting the observations. The focus of her observations would be the degree to which the students could answer oral and written questions after having read a passage from a book. Specifically, she would look for how her students used the strategies for reading comprehension that they had been taught.

Second, Kathleen also wanted to periodically ask her students direct questions regarding the use of those reading comprehension strategies. She designed a semistructured interview guide for conducting these student interviews. Her interview guide included the following questions:

- What does "reading comprehension" mean to you?
  - Do you have trouble understanding what you read?
  - Why do you think you have trouble?
- What helps someone understand what he or she has read?
  - Do you ever do any of these things?
- What strategies do you use to help you understand what you read?
  - Do you enjoy reading?

Kathleen planned to interview each student at least twice at roughly 2-month intervals during the course of her action research project. She anticipated learning more about their perceptions of reading, in general, and reading for understanding. She was also curious as to whether those perceptions would change over time.

RELATED WEB SITES: ADVICE AND GUIDELINES ABOUT DATA COLLECTION

The list of related Web sites for this chapter all come from the Action Research Web site of the Madison (Wisconsin) Metropolitan School District. They provide several good ideas and suggestions for issues you may very well face when dealing with decisions about collecting data for your action research project.

- **Guidelines for Data Collection** (http://www.madison.k12.wi.us/sod/car/cardataguidelines.html)

  Offered here is a bulleted list of suggestions, or simply things to consider, when planning for your data collection. The page begins with the following statement: "Asking the right questions is the key skill in effective data collection." Guidelines included on this list are "Be clear as to why you are collecting data," "Be clear about how you are going to use the data you collect," "Decide how much data is needed," and "Use multiple sources of data to increase the believability of the findings."

- **Techniques for Gathering Data** (http://www.madison.k12.wi.us/sod/car/cartechniques.html)

  Fourteen different techniques for collecting data are briefly described. The list includes many that we discussed in this chapter but also includes several additional techniques. These additional techniques include portfolios, still photography, and time-on-task analysis.

- **Data Collection: The Five Ws and an H** (http://www.madison.k12.wi.us/sod/car/car5wandh.html)

  The authors suggest that, prior to actually collecting data, teachers should ask themselves several questions, listed on this Web page. These questions fall under the following broader questions:
  Why are we collecting this data?
  What exactly are we collecting?
  Where are we going to collect data and for how long?
  When are we going to collect data and for how long?
  Who is going to collect the data?
  How will the data be collected and displayed?

SUMMARY

- Qualitative data are narrative, appearing primarily as words.
- Qualitative data are collected usually through observations, interviews, journals, or by obtaining existing documents or records.
- Observations involve carefully and systematically watching and recording what you see and hear in a given setting.
- Classroom observations may be structured, semistructured, or unstructured. Most teachers find it very difficult, if not impossible, to conduct structured observations of their classes.
- Unstructured or semistructured observations allow for the flexibility to attend to other events occurring in the classroom.
- Advantages of conducting observations include being able to collect actual data on student behaviors, which enables the researcher to see some things that students might not self-report. Videotapes can also help facilitate the observation process.

- Limitations of observations include that fact that making observational notes can change student behavior and that different observers can see very different things.
- Classroom observations are usually recorded in the form of fieldnotes, which may include observers' comments (which are preliminary interpretations of your observations).
- Interviews are typically formal conversations between individuals.
- Interviews typically follow an interview guide, which may be structured, semistructured, or open-ended.
- Structured interview guides provide a specific set of questions, which are asked uniformly of each individual being interviewed.
- Semistructured interview guides begin with several base questions but permit the researcher to ask follow-up, probing questions of individuals.
- Open-ended interviews ask very broad questions of participants, with the intent of gathering a wide variety of responses.
- Interviews can also be conducted with groups of individuals in an interview known as a focus group.
- Interviews may also be conducted informally or via e-mail.
- Journals may also be kept by both teachers and students in order to provide valuable information into the workings of a classroom.
- Existing documents and records, originally gathered for reasons other than action research, are abundantly available in schools and may be used as additional sources of information. These include classroom artifacts, such as student work.
- It is important for teacher-researchers to establish the trustworthiness of their data. This includes the accuracy, credibility, and dependability of one's qualitative data.
- Establishing the trustworthiness of qualitative data is accomplished through triangulation, member checking, and prolonged engagement and observation.
- Quantitative data are numerical and include just about anything that can be counted, tallied, or rated.
- Surveys are lists of statements or questions to which participants respond.
- Questionnaires are one specific type of survey involving the administration of questions or statements in written form.
- Items on surveys can consist of open-ended questions or closed-response rating scales.
- A closed-response question or statement provides the respondent with a number of choices from which to select. Analysis of the resulting data involves counting the number of responses for each option.
- Open-ended items allow for a seemingly limitless number of possible responses. Analysis of these data involve the categorization of responses into similar groups and then counting them.
- Scaled items can appear on surveys or rating scales as Likert scales (where individuals respond on an agree-disagree continuum) or Likert-type scales (where something other than extent of agreement is being measured).
- Surveys and rating scales are effective at gathering data simultaneously from numerous individuals but can sometimes be time consuming to analyze.
- Checklists are a simple form of rating scale where only a dichotomy of response options (e.g., present/not present) exists.
- Tests and other formal instruments can be used as quantitative data, provided they are supplemented with other forms of data.

- Validity of quantitative data have to do with the extent to which the data are what they are believed to be. It is best established through subjective analysis of the content of the items appearing on the survey, rating scale, or test to determine the degree to which it matches the underlying construct the researcher is trying to measure.
- Reliability refers to the consistency of quantitative data and is determined statistically.
- Remember the following: A valid test is always reliable, but a reliable test is not necessarily valid.

QUESTIONS AND ACTIVITIES

1. Describe what you might see as the benefits of collecting both qualitative and quantitative data as part of an action research study. Do you envision any potential negative aspects or weaknesses associated with collecting both as part of the same study?
2. Why is it important for researchers in general, and specifically for teacher-researchers, to take measures to ensure the quality of their collected data?
3. Making good, sound observations typically requires some training, or at least practice. Find a location with numerous people (e.g., a shopping mall, your student union) and spend 30 minutes observing and making fieldnotes on what you see and hear. Include any observers' comments as you deem appropriate during your period of observation. After the 30 minutes of observation, reflect on the experience. What did you think and how did you feel? How could you improve your observation and note-taking skills for a next observation session?
4. Think of a topic of interest to you and appropriate for an action research study. Develop a semistructured interview guide for a 15-minute interview with an individual. Be sure to include in your guide any "optional" probing questions. Next, interview someone using your guide. Afterward, reflect on your experience as an interviewer. What did you think and how did you feel? How could you improve your skills for a next interview?
5. Using the same topic you identified for number 3 above, develop a 15-item survey or rating scale that targets a specific audience, paying close attention to the guidelines presented in the chapter. Remember that a rating scale will use either a Likert or Likert-type scale. Administer your instrument to at least five individuals. Afterward, ask them to provide you with feedback on the instrument. Reflect on the process of instrument development and administration.
6. Can you think of any existing documents or records that would support an investigation of your topic? What are they? How difficult would it be to gain access to them?