CHAPTER FOUR

Knowing the Data Fundamentals

Data literacy can be inert—a skill for understanding and describing data—or active—a vital method for generating a knowledge source that informs action. With purposeful identification, collection, examination, and interpretation, data become energizing sources of information. Data options to explore extend beyond quantitative, wide-scaled measures, such as state achievement tests, and include a variety of both formative and summative sources. When groups collaboratively explore the data, rich conversations produce new understandings for group members about their students, their practice, their programs, and themselves.

Prodding, poking, and inquiring into what's going on, why it's going on, and whether it is satisfactory motivates change. A rigorous, data-driven process allows practitioners to better describe the current state of achievement and to identify gaps between the present state and their desired achievement outcomes. Potential gaps might exist between individual students' or groups' achievement and the high expectations we set for them, between a particular program's or curriculum's effectiveness and the standards we identify for successful implementation, and between one school's performance and another's. The collaborative learning cycle provides a framework for structuring productive, time-efficient, and balanced conversations.

Prodding, poking, and inquiring into what's going on, why it's going on, and whether it is satisfactory motivates change.

This chapter explains data fundamentals—data types, sources, terms, and tools—to support groups in effectively implementing the collaborative learning cycle and to broaden perspectives on the types of data that might further their work. Schools and school districts are rich in data. Determining how much and what types of data will best serve collaborative inquiry and the group's ultimate outcomes is critical to effective application of the collaborative learning cycle. It is important that the data involved provide both a broad and a deep view of the present picture but are not so complex that the process becomes overwhelming and unmanageable. By examining multiple sources of data, groups glean insights from several angles on the issues under study.

Two Types of Data: Qualitative and Quantitative

Fundamentally, there are two major types of data: (1) qualitative and (2) quantitative. Qualitative data rely on description, while quantitative data rely on numbers.

Qualitative data tend to be narrative, holistic, and longitudinal. Schools may gather qualitative data from the classroom, grade, department, or school level. Classroom-, grade level–, or department-based data include anecdotal records, student work samples, portfolios, student interviews, checklists, and homework assignments. School-level data include meeting agendas, teacher demographics (years of experience, education, and so on), memos, schedules, and curriculum maps.

Quantitative data are expressed numerically and statistically. Quantitative sources include test scores of all types, performance grades, attendance records, and enrollment data. Quantitative results are intended for comparisons between students, groups of students, schools, districts, states or provinces, and nations. Thus, they are expressed and described using stanines, quartiles, norm-curve equivalents, means, medians, and modes.

Each type of data is organized differently for analysis. Because quantitative data use numbers, percentiles, and other mathematical configurations, and they are organized based on frequency distributions, central tendencies, variabilities, and dispersions, it is easier to create tables, charts, and graphs to discuss and analyze quantitative data. For example, isolating, or disaggregating, norm-referenced test scores by variables such as gender, race, or socioeconomic status allows teams to identify relationships and patterns within the blur of numbers.

Qualitative data, which are descriptive, are usually reviewed holistically through examining anecdotes and artifacts. For comparative purposes and to discern patterns or trends, teams can organize these sources by frequency of instances, events, responses, products, or so on. Teams can use the categories or topics that emerge from the individual items identified for tables, charts, or further investigation. For example, by logging discipline issues by type, time of year, time of day, and school location, teams discover relationships from the flow of events that are often unseen by those too close to the individual occurrences.

It is important to note that while numbers are often used for scoring qualitative assessment, for example in rubrics and survey scales, that does not mean the data are quantitative. These numbers provide relative, and still often subjective, comparisons that rely on individual judgment. They do not provide data that are statistically comparable. Table 4.1 provides some examples of quantitative and qualitative data.

Table 4.1: Examples of Quantitative and Qualitative Data

	Quantitative Data	Qualitative Data
Student Performance Data	A variety of test results including proficiency tests, standardized tests, state exams, district, and classroom-based tests Report card grades The number of students receiving special services from local, state, or federal resources Attendance rates, mobility rates, expulsion rates, suspension rates, or dropout rates Percentage of high school graduates Percentage of students with disabilities who are mainstreamed into regular classes Percentage of retentions or advancements	Student portfolios and other work products Videotapes of student work and performances Exhibitions Student surveys, including pleasure-reading inventories, self-esteem stems, and self-assessment profiles Student journals and learning logs Observational records, anecdotal records, and running records Student interviews Checklists
Program Data	Teacher-student ratios Numbers of students enrolled in various programs like advanced placement Head Start, all-day kindergarten, and prevention and intervention programs Teacher and administrator education statistics, for example, education levels achieved, average number of years with district, average number of years of service overall, number of inductees (first three years), and number of retirements expected in the next three years Teacher participation in professional development activities Budget and resource allocations	Videotapes of special events, classrooms, and hallways Meeting agendas, minutes, and memos Teacher and administrator portfolios Artifacts like awards and photos of bulletin boards Staff interviews Workshop and training program agendas and evaluations Bulletins and newsletters
Community Data	Data on family demographics, for example, average income, percentage of single-parent families, and percentage of two-income households Number of school and business and industry partnerships Employment rate; employment sectors in the area	Focus-group data Opinion surveys Interviews with parents and community members

Source: Wellman & Lipton, 2004.

Formative and Summative Data

Both quantitative and qualitative data can be used formatively or summatively. Formative data illuminate what students know and can do and what they cannot yet do. Using data formatively provides feedback for teachers and learners to inform next steps and potential instructional modifications. Summative data are used to make judgments about student learning, effective instruction, or the effectiveness of a program or plan, generally at the end of an instructional period.

Formative data illuminate what students know and can do and what they cannot yet do. . . . Summative data are used to make judgments about student learning, effective instruction, or the effectiveness of a program or plan.

Formative Assessment

Dylan Wiliam (2011) offers a comprehensive definition of formative assessment that includes the work of several others in the field. Each description includes three fundamental aspects. Formative assessment (1) involves both teachers and students, (2) provides feedback for teachers to modify instructional activities, and (3) is applied during the instructional process for the purpose of improvement. Formative assessment is also process oriented. Various tools, instruments, and strategies are used to provide feedback based on the interaction between teacher and student, student and student, and student and information or learning task. It is active, responsive, and adaptable. The information provided is just as important for the students as it is for the teachers, if not more so.

When teachers collaboratively explore formative assessment data, they can look for patterns and determine adjustments in teaching to improve learning. Through regular, ongoing monitoring and then focusing interventions and adjustments accordingly, teachers optimize student learning and increase student performance on summative measures that determine judgments about performance, programs, and practitioners.

Summative Assessment

Teachers collect summative data in short cycles at the end of instruction, as in benchmark data, or at the end of a course. Summative assessment can be quantitative, such as unit tests and state or provincial exams, or qualitative, such as rubric-driven projects or student portfolios.

It is best to use summative data to make decisions about curriculum, to direct future instruction, and to improve professional practice. Often, teachers use yearly summative data to begin a collaborative learning cycle and apply identified formative sources at appropriate intervals for continuous improvement.

Reliability, Validity, and Credibility

Here we offer definitions and distinctions between several terms related to data. While *reliability* and *validity* tend to be terms applied to quantitative data,

the concept of *trustworthiness*, or *dependability*, is still important when considering data sources and results. The concern for confidence in the findings brings the idea of *credibility* to the forefront. Whether or not results are considered reliable, valid, trustworthy, or dependable is often subjective and is more about the viewer than the viewed.

Reliability

Reliability refers to consistent measures or results from a particular tool or instrument. An instrument is considered reliable if it produces the same result time after time. For example, if a test is designed to measure computation skills for first graders, then each time the test is administered to first graders it should produce essentially the same results. The concern for reliability is one reason that schools tend to rely heavily on norm- and criterion-referenced tests. These measures are generally reliable across time and for a wide range of students. However, the test items are limited regarding what they can measure, as they are designed to fit a neutral, objective machine-scoring system. Open-ended questions—those requiring responses that show inventiveness, creativity, or examples intended to reveal the process of problem solving—do not fit these systems. While more authentic measures, such as performance tasks and portfolios, can provide data on these skills and attributes, they raise the problem of *inter-rater reliability*. This problem occurs when different scorers arrive at different results, or when an individual scorer's ratings over time show wide differentiation. Failure to invest time and energy in developing shared performance standards and scoring accuracy reduces the effectiveness and credibility of performance tasks. Scoring rubrics, exemplars, and training sessions designed to increase inter-rater reliability are intended to reduce this concern.

Validity

While reliability is concerned with the accuracy of the instrument based on the consistency of results, validity refers to the degree to which an instrument measures what it is designed to measure—that the data reflect what they are intended to show. For example, many fill-in and multiple-choice tests are intended to measure content knowledge, but may be more likely to measure a student's syntactical knowledge or reading ability. One subset of validity regards content. *Content validity* indicates how comprehensively an instrument measures or completely tests for the domain of knowledge it is intended to assess. For example, using an assessment that only includes items that test decoding to determine comprehension skills would indicate a lack of content validity as the broad range of comprehension skills would not be measured.

Further, although standardized, norm-referenced tests are highly reliable, they may not be valid in measuring sophisticated conceptual understanding, or the ability to solve complex problems. Such tests may also contain questions that permit students to be credited with correct answers for the wrong reason, give away the answer in the options presented in multiple-choice questions, or

contain language that is linguistically or culturally confusing to students. (For insight into this issue and specific examples, visit http://marciakastner.com/10-common-flaws-in-math-tests_289.html; Kastner, 2009a).

Credibility

An additional concern when considering what data to explore is that of credibility. An individual's experiences, knowledge, and confidence using different data tools will influence his or her perception of its believability, or credibility. No matter how reliable and valid an instrument may be, differences in learning style and educational philosophy, as well as diversity in gender, ethnicity, professional position, and so on, influence a group member's approach to and participation in data-driven inquiry. By incorporating multiple data types and multiple sources in a data exploration—qualitative and quantitative, classroom-centered, and systems-based assessments—leaders will ensure that more group members can relate to more of the data. In other words, the data are more credible because they are within the scope of individual group member's readiness to accept them.

Triangulation and Disaggregation

Both triangulation and disaggregation give data teams clearer and more precise information from which to work. *Triangulation* is an effective method for increasing credibility and dependability when in a data exploration. Imagine that an event occurs in a central plaza between three buildings. In each building, there is someone at a window watching the scene unfold. No single perspective could offer a full picture of what occurred. To get a comprehensive description, information from each of the three observers is needed. When exploring data, the same concept applies. Triangulation requires multiple views, or multiple sources, to create a more complete assessment of the issue being studied.

Disaggregation, or breaking a large swatch of data into smaller subsets, often reveals what may remain hidden or buried. Discerning the part from the whole gives a more precise view of those parts and often reveals more information about the underlying issues within the larger whole.

Using Multiple Sources: Triangulation

No individual assessment or measurement instrument is a perfect fit for providing what we want to know about whom and in what ways. Using multiple data sources compensates for the deficits in individual tools and provides a comprehensive picture of the topic under study. Together, these data sources provide a more comprehensive picture than any individual data set could. *Triangulation* is a researcher's term for taking at least three different perspectives on an area of study.

One view alone offers a limited and usually too narrow viewpoint. For example, standardized test scores for School A indicate that 64 percent of the

students in grade 8 scored proficient or above on the state language arts exam. A performance ranking of the eighth-grade students reveals a clearer understanding of where within each band these students performed. Demographic data on each of the students offer even more information about groups that may be marginalized, or specific cohorts whose learning needs are not being met. Triangulation is one way of addressing validity. We can reasonably assume that if three different measures are all indicating similar results, they offer valid information regarding what we want to know.

Triangulating the data is most powerful when the various sources are diverse and varied. Thus, a qualitative measure such as a student survey will enhance quantitative measures like state or provincial exams. A third source that is different from either of these, such as teacher anecdotes or a curriculum map, can enhance the view. The idea is to seek multiple sources, using different methods and operating at different levels or in different areas of the school. This approach also addresses credibility in that multiple sources of differing data types are most likely to be credible to different group members. Intentionally seeking multiple perspectives, both of the data and those analyzing them, enriches the process and outcomes.

Making the Invisible Visible: Disaggregation

Disaggregation is breaking the data apart and reorganizing it into smaller subsets. These data might be subscores within a larger measure. For example, disaggregated scores are often provided for criterion- and norm-referenced tests by skill strands. In addition to the total reading score, the results provide a look at literal and inferential comprehension or vocabulary skills. Or data might be disaggregated based on a specific characteristic, such as viewing math scores by gender or exploring for correlations between reading performance and specific socioeconomic groups.

Disaggregation addresses important questions about what is working (or not) and for whom. For example, disaggregating data by gender helps determine whether an improvement in math achievement reflects equal gains for male and female students. The same questions can be explored for any subset to determine and ensure that all students have equal access and opportunity to learn.

Disaggregated data give much clearer, more specific information than holistic, blanket scores. Keep in mind that there are subsets within subsets. It is important to sort a variety of variables, but not so many that the data become overwhelmingly complex or that the subgroup becomes too small.

Data that are disaggregated by race or ethnicity, gender, income, academic programs, geographic area, feeder schools, and even classroom teachers will often reveal patterns and provide indicators of present and potential future performance that remain hidden in large data lumps.

Dimensions of Data

To choose data is also to choose perspective. Just as the vista from the top of the mountain is quite different than the view from the valley, varied data offer varied lenses on the world of learning.

To choose data is also to choose perspective. . . . Varied data offer varied lenses on the world of learning.

Information Altitudes

One important dimension of data is its distance in time or level from the topic under study. Different types of data offer wide-angle or telephoto images of student performance. For example, to explore middle school students' achievement in science, the data might include individual student scores, grade-level scores, schoolwide scores, and perhaps districtwide scores. Each set of scores might include normed indices, such as the Northwest Evaluation Association Measures of Academic Progress (MAP) or the Trends in International Mathematics and Science Study (TIMSS), provincial- or state-level scores measuring provincial or state standards, as well as district, school, department, or classroom-based tests. In addition, data might include student work samples, homework assignments, projects, journals, or portfolios. We think of these measures based on their distance from the initial inquiry and the dimensions of the field of focus. If investigating the question, How are our students performing in science at the middle school level?, the individual student work would offer the closest view, while the nationally or internationally referenced tests, the farthest. (See figure 4.1 for an example.)

Temporal Dimension

Data also have a temporal dimension. That is, we can explore data that have been archived from past events, we can observe and collect present samples, and we can project measurements and tools that will yield information into the future. Once again, if we're interested in exploring the effectiveness of our science instruction, we can look at standardized achievement scores for the previous three years to surface trend data (past). We can also conduct surveys regarding students' attitudes toward science and demographic data about the science teachers (present). In addition, we can identify data that will be useful when they become available, such as scores on end-of-grade tests or state science exams (future).

Specific Questions

A question-driven approach is a powerful way to motivate and mobilize energy for school improvement. Questions related to differences, similarities, gaps, qualities, characteristics (both desired and existing), patterns of success, impact of programs, curriculum, instructional methods, and so on provide relevant starting points for a data pursuit. This approach is true whether the questions emerge from a data exploration or provide the impetus to begin one.

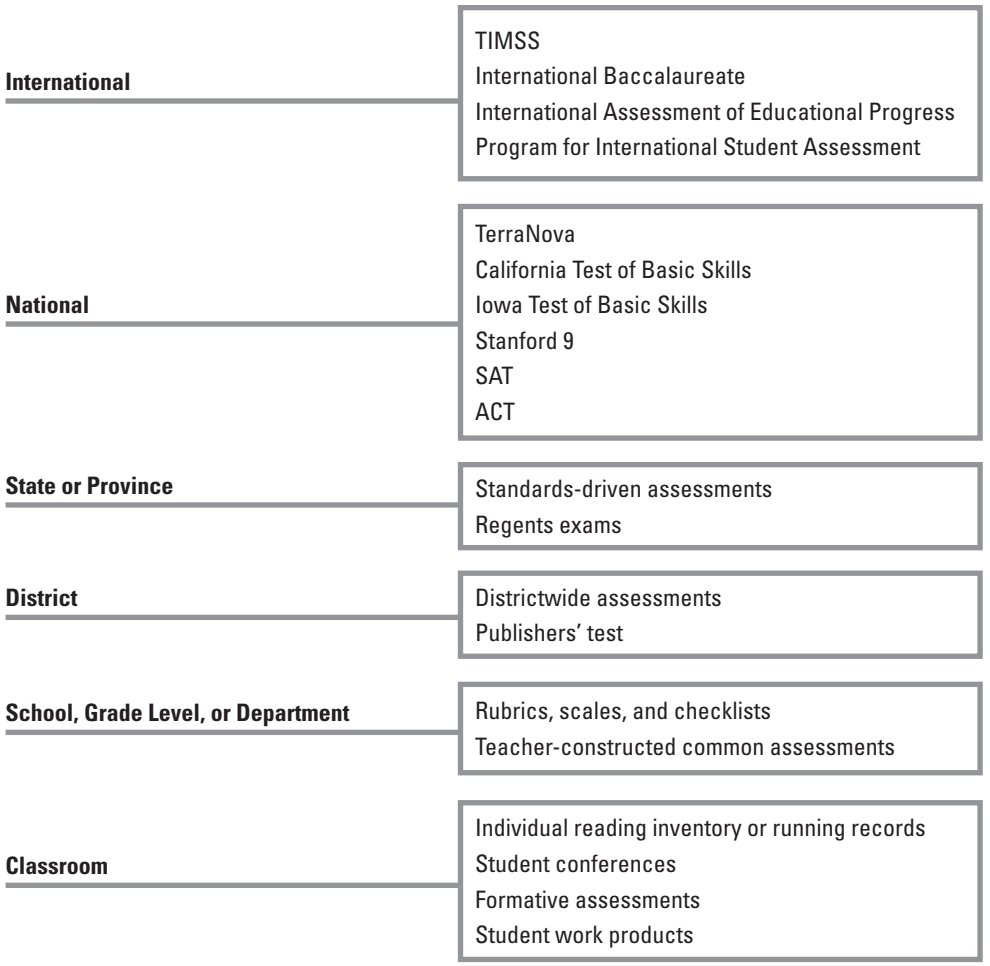


Figure 4.1: Information altitudes.

As described in chapter 3, sometimes the inquiry process begins with a question, such as, "In what ways does use of technology in the classroom affect students' research skills?," or a problem statement, such as, "Sixth-grade students enter middle school without the fundamental skills needed for proficiency in math." Or the inquiry process might begin with a hypothesis, such as, "Incidences of bullying would decrease if the curriculum included more social and emotional outcomes." Each approach still requires the identification of the data needed to explore the issue.

Another decision point involves the scope of the topic being explored. Herman and Winters (1992) describe the options as *wide angle* and *close-up*. Wide-angle investigation, whether it is at a district, school, department, or grade level, is often a good beginning for study. Close-up investigations narrow and hone the focus. For example, a district-level group might want to examine the big picture, "How effective are our science curriculum and instruction?," while a school-based team

might be most interested in the question, "How has incorporating nonfiction into science units paid off for third- to fifth-grade learners?"

Wide-angle questions concern the programs' effectiveness for all learners, or large subsets. Examples of wide-angle questions include inquiry into achievement gaps among different groups in particular content areas ("Is there a positive correlation between students' skills and their performance in math? Is the rate of enrollment in advanced placement courses for minorities different than that of other groups?") or into major issues for the school population ("Is student transiency affecting graduation rates? In what ways are temporary teacher certifications affecting school discipline?").

Once a data team frames large issues and designs approaches for improvement, exploring close-up questions offers additional insights and opportunities to monitor the plan's effectiveness. Often these are the formative questions that lead to refinement and enhancement of an improvement plan. Examples in this category include questions about new methods in assessment ("How well do our new performance tasks measure our elementary students' achievement in science?") or instruction ("In what ways has differentiation improved performance for special needs students?").

The inquiries should be clear and contextual, and they should create conscious curiosity—not fear of reprisal or evaluation.

Most importantly, the inquiries should be clear and contextual, and they should create conscious curiosity—not fear of reprisal or evaluation. Clear questions emerge from data explorations and collaborative focus on critical issues. These inquiries often illuminate the most useful data for shedding light on areas of concern. These data choices complement and extend existing data, amplifying understanding. For example, to explore the effect of a new districtwide program on students' performance in math, we would access math achievement scores and compare the present year with the previous ones found in archival data. Data on teachers' instructional repertoire and staff participation and training outcomes for related professional development workshops would also be useful. In addition, a survey of students' perceived confidence in math or math teachers' confidence in teaching might add useful information. The workshop sign-in sheets and student and teacher surveys are examples of collectible data.

Collectible Data Tools

Additional collection tools include checklists; anecdotal records, such as running records and teacher observation logs; videos; and charts, such as seating charts, agendas, and meeting minutes.

Collectible data usually sort into two categories: (1) behavioral and (2) perceptual. *Behavioral data*, or observational data, include the workshop sign-in sheets described previously. Other examples include frequency of office referrals or percentages of increase or decrease in parent attendance at school functions. Checklists and anecdotal records are useful instruments for collecting these data.

Perceptual data include interviews with various groups regarding their feelings, understandings, or opinions about a particular issue, innovation, or program. Opinions on the degree of satisfaction with the new block schedule, perceptions regarding effectiveness of the present reading program, and parents' responses to homework policy are additional examples of perceptual data. Interviews and surveys are useful tools for collecting these types of data.

Interviews and Surveys

Protocols, such as interviews and surveys, formalize the inquiry process as instruments for data collection. The following tools are examples of commonly used methods for this purpose. An interview is essentially a survey conducted either face-to-face or via telephone. Survey instruments are generally distributed and returned electronically or through the mail, although they can be completed at a meeting. It is important to note that self-reports provide powerful perceptual data, but they are also limited in accuracy and should be triangulated with other methods.

Interviews

Interviews are a source of qualitative, perceptual data. The intention is to record as fully and fairly as possible each respondent's particular perspective. Interpretation and analysis occur after interview data are organized—never during the interview process. It is important, therefore, to maintain a stance of inquiry when collecting interview data. Interviews vary in their degree of formality. The most informal interviews are conversational and responsive, the questions emerging from the interaction. The most formal interviews are based on carefully constructed, and often field-tested, protocols. Table 4.2 illustrates three types of interviews, ranging from least to most formal.

Table 4.2: Three Types of Interviews

Interview Style	Description	Suggested Uses	Cautions
Informal Conversational Interview	Relies on spontaneously generating questions in the natural flow of the interaction	To explore beliefs, attitudes, values, or perceptions	Requires a high level of inquiry skills
General Guided Interview	Based on an outlined set of issues established in advance, which serves as a basic checklist to ensure all relevant topics are explored	To explore complex issues that do not have finite or predetermined responses	Requires thorough preparation May be time consuming both in preparation and implementation
Standardized Open-Ended Interview	Consists of a carefully constructed protocol Same sequence and wording repeats with each interviewee	To increase generalizability of results To increase reliability and validity of research	Requires a high level of consistency May require interviewer training

Source: Adapted from Wellman & Lipton, 2004.

Surveys

As with interviews, surveys also provide an effective method for collecting both behavioral and perceptual information. Researchers most commonly use two primary survey categories: (1) scaled and (2) unscaled.

1. **Scaled surveys:** A scaled survey asks respondents to quantify their answers and provides information that is readily organized graphically. However, these responses are often limited in scope. For example, a response of 4 to the question, "On a scale of 1–5 how would you rate the effectiveness of this program?," does not offer information on what constitutes a 4 for this respondent, or what might have made it a 5.
2. **Unscaled survey:** Unscaled surveys usually require responses to open-ended questions. Compared to scaled surveys, these instruments yield more potential information, but are more complex to organize for analysis. For example, a question might read, "In what ways did this program meet your expectations?" These qualitative data are most often organized into tables or charts based on emerging categories.

Creative Research Systems (2011) offers five basic steps to survey design.

1. **Establish the goals of the project:** Determine what you want to learn or know more about.
2. **Identify your sample:** Determine who will participate in the survey.
3. **Choose survey methods:** Design how you will gather the information (interview or written response).
4. **Create your questions:** Decide what you will ask and in what ways. See Types of Survey Questions (page 72) for examples. (Visit **go.solution-tree.com/teams** to download the reproducibles in this book.)
5. **Pretest or pilot the survey:** This optional step, if practiced logistically, is useful for testing the questions, protocol, or both.

Feasibility

Feasibility is the term for how realistic and reasonable it is to collect and apply a data source. The realities of life in schools create a critical need for making effective and efficient choices. Time, money, energy, and even space come into consideration in making choices about what data to collect. One tip is to look first to archival data, those things that teachers, schools, and districts are already collecting and have readily available. After tapping archival resources, groups can make creative and selective choices identifying additional collectibles that will illuminate, detail, and complement the existing data.

Archival Data

Archival data are those that already exist and are accessible, often electronically, as part of the district's established information base. These can be both quantitative and qualitative sources. Archival data include student performance

data—such as test results, grades, referral, suspension, and retention rates—and demographic data—such as percentage of students in particular programs; race, ethnicity, and gender profiles; attendance rates; and socioeconomic status (for example, numbers of students receiving free or reduced lunch). Demographic data for staff might include years of experience, types of certification, and levels of education. Mobility rates can be accessed for students, staff, and administration.

Much archival data have an anthropological quality and include sources that offer insight into the day-to-day life in the district, school, or classrooms. These data include correspondence such as memos, newsletters, bulletins, and meeting minutes; descriptions of course offerings; and available extracurricular activities. Per-pupil expenditures and other budget information fit this category, as well.

Because of its longitudinal quality, archival data can be accessed and analyzed for trends. For example, a study group might create a line graph comparing student reading scores over five years or chart the number of students participating in various extracurricular activities over a three-year period.

In addition, school archives offer access to data that relate to specific inquiries or improvement efforts. Let's say that a school improvement plan has the goal of increasing mathematics performance. Important data sources might include curriculum maps or lesson plans to determine a baseline on how much time is spent teaching math. This information can be disaggregated or subdivided by grade level and percentage of teachers trained in the current math program or have specializations in that area and inventories of instructional materials and resources (for example, which textbooks, kits, and other published materials teachers use).

Visually Effective Data Displays

Well-crafted data displays clarify and communicate often complex or abstract information. For groups exploring data, visually vibrant displays capture and focus attention. As discussed in chapter 2 (page 32), a large data display serves as a third point, physically separating the group from the data and objectifying the data—making data a *thing*. As a result, the conversation is about what we (colleagues) notice and think about them (the data). This combination increases the emotional safety needed for group members to poke, prod, and question the data and one another.

Well-crafted data displays clarify and communicate often complex or abstract information.

In contrast, distributing individual and often cluttered printouts dissolves group cohesion. Participants drop into personal searches, viewing the data from their own vantage points. As a result, group members may be sitting side by side, but lose access to one another's perspectives and experiences. This lack of shared interaction limits shared understanding and collective commitment to action.

Effective data displays should be clear, vibrant, and adequately sized for data groups to share them. A ratio of approximately four to six participants to one large data display works to create a focal point. The small group size increases participation and captures individual energy.

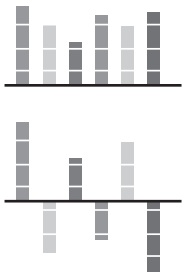
Making a display is not a neutral act. The purposeful consideration and choice of display make different data types more accessible to group work. A particular display shapes the conversation by illuminating relationships in the data set and inviting different kinds of inquiry and analysis.

Choosing the Visual Representation

Effective collaborative inquiry requires a recasting of raw data or typically used tables and charts into more compelling, clear, and accessible displays. In too many cases, displays in schools and school districts are crowded, unclear, and difficult to read. This problem is exemplified by many state and provincial reports.

Following are six common errors in displaying data.

1. **The display does not illustrate relationships within the data:** For example, using a bar graph to illustrate a trend. To show a trend over time, a line graph would be more effective. Choosing the most effective graphic format facilitates the work of data teams. (See figure 4.2.)
2. **The display range visually skews the values:** For example, plotting performance scores ranging between 35 and 72 percent on a vertical axis of 30 to 80 percent. Plotting the scores on a broader scale from 0 to 100 percent would more accurately reflect results.
3. **The display uses colors with no or low contrast:** For example, using several shades of blue for different subgroups or trend lines. Use contrasting colors for different bars, lines, or pie slices to eliminate visual confusion. Note: Save really bright colors for information that needs to stand out.
4. **The display attempts to squeeze too much information into one chart or graph:** For example, including several subject areas, multiple grade levels, and subgroups in one display. Create several different displays to make the data less overwhelming.
5. **The display labels are difficult to read:** For example, labeling the horizontal (or *x*) axis with vertical type or using a font that is too small is hard to read from a distance. Create labels that are clear and easy to read. Note: Type set on an angle is also ineffective.
6. **The display uses graphic effects without purpose:** For example, overusing 3-D effects. Flashy displays create visual clutter and can actually hide some of the values when bars overlap. Leave the special effects to video games.



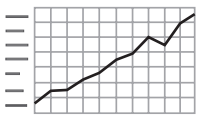
Bar Graphs

Bar graphs can be used to display comparisons, rankings, and change over time, for example, performance in two content areas or performance by subgroup or proficiency levels. Stacked bars show the elements that comprise the total while distinguishing the relative size of the parts, such as stacking proficiency levels within a whole grade level.

Deviation bar graphs display the data above and below a baseline, or the pluses and minuses of an issue or deviations from a standard, such as a locally established benchmark or a nationally referenced norm.

Pictographs

Representing data as pictures makes a creative and eye-catching variation from the more commonly used graphic elements. For example, pictographs can replace the icons on a line graph or can be stacked to form the bars in any form of a bar graph.

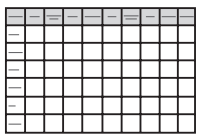


Line Graphs

Line graphs display a sloping line or segments of lines, representing change over time. Thus, they are particularly useful for displaying trends, such as reading scores over a period of years. Line graphs can also display comparisons when several lines are used on the same graph, such as reading scores for several schools or several grade levels over several years.

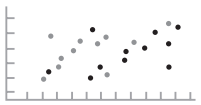
Pie Charts

Pie charts display parts of the whole. The size of each part displayed as a percentage makes the relationships among the parts and between the part and whole graphically apparent to observers. For example, pie charts can be used for demographic data, such as teachers' years of experience or students' socioeconomic status.



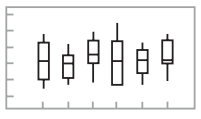
Tables

Tables show exact numerical values, such as scores for each student in a class. They are used to portray simple data sets that compare related values, such as scores by teacher by year. Tables are also effective for displaying quantitative information involving more than one unit of measure, such as students by age by performance quartile.



Scatter Plots

Scatter plots display relationships between two or more variables. They indicate correlations and comparisons, at one point in time or over time, such as scores in reading comprehension and math problem solving.



Box and Whisker Plots

Box and whisker plots turn raw data into the "shape" of the score distribution, such as standardized test scores, for ease of visual interpretation. The boxes display the distribution of scores, while the whiskers indicate the range above and below the median.

Figure 4.2: Graphic formats for visual data display.

Visual representation is best based on the data's function. . . . The display you choose should be closely related to the question, issue, or topic being explored.

Visual representation is best based on the data's function. In most cases, data displays are developed from raw data or, more often, data tables. The display you choose should be closely related to the question, issue, or topic being explored. For data sets of twenty numbers or less, tables are more effective than graphs, as they offer more specific information. Tables and charts, including frequency charts, disaggregation tables, and curriculum maps, can be organized as useful displays. Other visuals, such as timelines, schedules, calendars, and other artifacts (student work samples, meeting minutes, sample memos, or newsletters), can also be enlarged to enhance a data display.

Stephen Few (2004) describes four common ways of navigating or interacting with data that inform display choices. Effective displays:

1. Filter out what's not relevant
2. Sort the data to see it in order of magnitude
3. Provide high-level (big picture) and low-level (detailed) views of the data
4. Offer varying views and perspectives of the data using different display types

Well-designed displays reduce the mental effort required to interpret the data, allowing the group to focus on exploring and making meaning from the data.

Data Story: Using Triangulation

For a full year, New City School District has been implementing a three-pronged plan to improve math performance: a revised curriculum, a new instructional program, and a cohort of building-level math coaches.

At Northside Elementary School, the math coach is leading a vertical team of six teachers through the collaborative learning cycle. They are meeting to generate causal theories and determine additional data to guide their plans.

The coach begins.

> "So, based on our state test, problem solving is jumping out as an issue for all grade levels."

Group members chime in.

> "Well, I know at fourth grade, it's because students don't come with good computational skills, and that takes them off track. So, we need to look at whether that might be an effect across all grade levels."
>
> "Well, for my kids, it's more about their reading comprehension. They don't know how to read the problem to sort for critical information."

The coach intervenes, with a process suggestion:

> "Before we get too far, let's step back a bit and see what patterns appear in these state data."

The group breaks into pairs and reviews the data displays. It appears that there is not a significant, schoolwide gap in computation skills. Members agree to take a closer look at the issue.

Teachers decide to review the data, class by class and kid by kid, comparing the state math data to the state reading scores for three years. The strongest pattern that emerges across this period of time is the correlation between reading comprehension and problem solving, with gaps in both areas for a large percentage of students. Computation does not seem to be an issue overall.

The coach continues, "Given this additional perspective, what are some causal theories we might pursue?"

> "Well, I think the problem is that kids who aren't strong readers may have computation skills, but can't apply them inside of a word problem. They have difficulty interpreting the question, so they don't know which math operations to use."
>
> "If that's true, then we would want to look at using math word problems in our reading comprehension lessons to increase their skill in understanding what a problem is asking for."
>
> "Before we start overhauling our lesson plans, what can we use to confirm that we've nailed the problem? I think we need to get some grade-level and classroom-based data."
>
> "Yes, we need to see whether the same pattern holds for the assessments built into our new curriculum."
>
> "We can also script some think-alouds from kids at both ends of the performance range at each grade level."
>
> "We should also look at which grades and classrooms are getting the best results in this area, so we can learn more about what's going on."

The coach suggests that the group members begin by collecting think-alouds from their own students and recruiting other grade-level teachers to do the same. She will organize the curricular assessments and provide data displays for the next meeting. The coach brings the meeting to a close by charting these agreements and clarifying logistics for the next meeting.

Exercise Your Learning

Use the information from this chapter to do the following:

1. Design a data collection tool that you might use to gather data on an issue you are presently exploring.

2. Complete the Information Altitudes Exercise reproducible using a specific skill set.
3. With samples of your own data, create a visually vibrant display. Consider multiple options, and be prepared to explain your choices.

Extend Your Learning

Perceptual Edge (www.perceptualedge.com) is a source for learning how to design simple information displays for effective analysis and communication by Stephen Few. Other valuable sources include *Now You See It: Simple Visualization Techniques for Quantitative Analysis* (Few, 2009) and *Show Me the Numbers: Designing Tables and Graphs to Enlighten* (Few, 2004).

Edward Tufte (www.edwardtufte.com/tufte) is a source for increasing the sophistication and visual intelligence of your graphic displays and to learn how to turn data into engaging information. Edward Tufte (1983) is the author of *The Visual Display of Quantitative Information*.

Visit http://marciakastner.com for insights on the impact of poorly designed test questions.

Information Altitudes Exercise

Think of a specific skill set or knowledge base that you want to explore, such as reading comprehension or math problem solving. Using the following table, identify data sources at each level of information, as indicated.

Area of exploration: _____

Information Altitudes	Data Source
International	
National	
State or Province	
District	
School, Grade Level, or Department	
Classroom	

Types of Survey Questions

Background and Demographic Questions

- Concern the interviewee's identifying characteristics
- Are distinguished by their specific and somewhat routine nature

  Examples: How many years have you been teaching here? What level of education have you completed?

Experience and Behavior Questions

- Ask about what a person does or has done
- Elicit descriptions of observable experiences, behaviors, activities, and actions

  Example: If I had been in the classroom with you, what might I have seen or heard?

Knowledge Questions

- Designed to find out about the respondent's perspective on empirical data
- Seek factual information such as rules, regulations, program data, logistics, and so on

  Examples: How many planning periods are scheduled each week? How many students are scheduled for special services?

Cognitive Questions

- Aimed at understanding the cognitive and interpretive processes of people
- Tell us about people's interpretations, analyses, and inferences

  Example: What do you think about _____?

Affective Questions

- Seek to understand people's emotional responses to their experiences
- Will often be responded to with adjectives (such as *happy*, *responsible*, *intimidated*, or *frustrated*)

  Examples: How do you feel about _____? What are your reactions to _____?

Sensory Questions

- Ask about what is seen, heard, touched, tasted, and smelled
- Elicit vivid descriptions or events or environments

  Example: *When you walk into the classroom, what do you see?*

Identity and Value Questions

- Seek expressions of identity, values, and beliefs
- Tell us about interviewee's goals, intentions, desires, and values

  Examples: What do you value most about _____? What would you most like to see happen?

Role Play and Simulations

- Provide a context for potentially difficult questions
- Put the interviewee in the stance of expert
- Reduce the personal nature of some questions

  Examples: Suppose I just started teaching here, what would be the most important things I would need to know to be successful? Suppose you needed to get something done around here, what would you do?