

1

INTRODUCTION TO MEASUREMENT

INTRODUCTION

In this chapter we will begin to look at the various ways in which we use measurement in the classroom. We will differentiate between the terms assessment, measurement, and evaluation as well as differentiate between formal and informal assessment. We will look at the various roles of assessment including preliminary, diagnostic, formative, and summative assessment. We will also differentiate between maximum and typical performance measures. Finally, we examine a brief overview of the remaining chapters on measurement.

THE ROLE OF MEASUREMENT

Measurement plays many different roles in our lives. During the past week I have had to rely on measurement in many different ways. The other day, while at the grocery store, I had to locate a can of diced tomatoes that was 14.5 ounces for a recipe for Spanish rice. Each morning I rely on my alarm clock to wake me at 6:45 a.m. (I get to sleep-in late this year!) I have recently anxiously watched the price for a gallon of gasoline rise and fall. When a student asked me for a recommendation, I reviewed his performance in my class by checking my grade book, paying special attention to his test scores. These examples represent only a few of the many ways that I use measurement on a daily basis.

You will also use measurement a great deal in your classrooms. You may give students diagnostic tests to assess how well they are reading, spelling, writing, counting, and so on. You will give tests to assess how well your students have learned and mastered the material that you have been teaching. You will use informal assessment on a daily basis to judge how well your lessons are progressing: Do your students appear to be comprehending your instruction? You will collect homework to assess if your students are actually able to apply the new skills that they have been learning. As you will see, measurement will play many different roles in our classrooms.

In the five chapters of Part II we discuss statistics. That discussion is designed primarily to give you sufficient background to be able to examine and comprehend the many

measurement concepts that we will be discussing in the remaining 17 chapters. Those statistical concepts that will be most often used in the discussions of measurement include the mean, the standard deviation, the variance, correlation, and z-scores. You will see how we use these and other statistical concepts to build good measurement practices.

MEASUREMENT, ASSESSMENT, AND EVALUATION

There are many terms that describe the processes that we use to judge student performance. The three terms most commonly used are measurement, assessment, and evaluation. Although I will provide you with specific definitions for each term, I must warn you that there is no international law or agreement on how these terms should be used. I have seen many instances where these three terms are freely used interchangeably as if they all meant exactly the same thing. However, I will give you specific definitions for each according to the ways that each term is used most frequently in education.

Assessment

The most general term is assessment. *Assessment* is a very general term that describes the many techniques that we use to measure and judge student behavior and performance. Although in some professional circles the term assessment sometimes means something more specific, in relation to the classroom, it is typically a very general, very generic term. When in doubt about how to label a technique, the safest label to use would be assessment.

Measurement

Over the years, I have seen many definitions of measurement. However, here is the one that I find the most useful. *Measurement* is the process of assigning meaningful numbers (or labels) to persons or objects based on the degree to which they possess some characteristic.

Let's look at some examples and start with how we could take measurements of an object. What are some characteristics of a classroom to which we could assign meaningful numbers or labels? The first thing that we generally think of is the classroom's dimensions, especially its length and width. A second important characteristic is the classroom's seating capacity. Some classrooms may be designed to hold 20 students, whereas others could accommodate 35 to 40 students. This becomes important when we are assigning classes to certain rooms. Still another characteristic that I find important is board space. I tend to use the board a great deal, especially in some classes, and want classrooms with plenty of board space (perhaps measured in square feet). What about the room number? For example, let's say that we are in room 231. Is the room number a form of measurement? It may be! The room number is frequently used as a substitute for a name and can be considered measurement at the nominal (naming) level (see Part II, Chapter 18). If room numbers are assigned to rooms randomly and don't assist a stranger in finding the room, then I would argue that in that case the room number was not a form of measurement. In that case the room number was not very meaningful. However, if room numbers are assigned in some logical order (as they are in most buildings) and they assist a stranger in finding the room, then I would argue that they are measurement, if only at its least sophisticated level. In this case the room

number has some meaning and tells you something about the classroom, where it is located.

For another example, let's use a student. What are some characteristics of a student that can be described with meaningful numbers or labels? Of course, many of the student's various physical characteristics such as height, weight, and age can all be described with meaningful numbers. Even describing other characteristics such as gender are frequently considered meaningful at the nominal level. Other characteristics to which we can ascribe meaningful numbers would include the student's score on the last social studies quiz, the student's score on an I.Q. test, or her score on the standardized math achievement test. There are many student characteristics that we can measure in a meaningful way.

Measurement is primarily a mechanical process. When we score a test, count up the number of points that the student earned, and record that number, we are using measurement.

Evaluation

The third term is evaluation. *Evaluation* involves the use of measurement to make decisions about or to determine the worth of a person or object. Evaluation is frequently the step that follows measurement. Let's say that Dana earned 87 on a 100-point science test. When we scored Dana's test and determined that she earned 87 out of 100 points we were using measurement. It was a mechanical process. If we then decide that Dana's test score of 87 translates to a letter grade of B, we are now using evaluation. Deciding on a grade was dependent on a judgment of worth. A letter grade is typically a statement of worth: A is excellent, B is good, C is average, and so on. So, scoring a test is measurement, but assigning a grade is evaluation.

Evaluation comes in many forms. One obvious form is the use of letter grades. However, we sometimes run courses on a pass/fail basis. If your mean test score is 75% or higher, you pass. If it is below 75% you fail. At other times teachers may prepare a narrative statement concerning how well the student is performing and use terms like "making satisfactory progress," or "still needs more work in this area." These are also forms of evaluation.

Definitions

Assessment is a very general term that describes the many techniques that we use to measure and judge student behavior and performance.

Measurement is the process of assigning meaningful numbers (or labels) to persons or objects based on the degree to which they possess some characteristic.

Evaluation involves the use of measurement to make decisions about or to determine the worth of a person or object.

There are several measurement issues that it would be helpful to address in this chapter. First, we will differentiate between formal and informal assessment. Then we will look at the various ways in which we use assessment in the classroom. Finally, we will differentiate between maximum and typical performance measures.

FORMAL vs. INFORMAL ASSESSMENT

Generally, when you think about assessment in the classroom, I would guess that tests and quizzes are probably the first things that come to mind. Tests (both standardized and teacher-made), quizzes, and other similar devices are referred to as *formal* assessment devices. With formal assessment you, the teacher, are able to complete the assessment in a relatively standardized manner and are able to control many aspects of the process. With formal assessment all students are given the same questions to answer, typically have exactly the same amount of time to respond, and complete the device in a relatively consistent fashion. Formal assessment is important in the classroom; it is not, however, the only type of assessment that we use.

A much more common type of assessment is what we refer to as *informal* assessment. It involves the many observations that we make about students and the many questions that we ask students throughout the day. It may involve having a child read aloud and attempting to note and remember which reading processes are still giving this child problems. It could involve having children go to the board to work out math problems and looking for which students are performing well and which students appear confused. It can involve the many questions we ask students throughout the day and noting how each student responded to those questions. Many times we are using informal assessment to determine if our students are displaying competency with the skills which we have been teaching. Are we ready to move on to the next topic? Clearly, we do a great deal of informal assessment in the classroom. In fact, I have heard educators estimate that perhaps as much as 90% of the assessment that we do in the classroom is informal assessment, especially in the early grades.

CLASSROOM ASSESSMENT

Preliminary or Placement Assessment

We use assessment in the classroom to serve a variety of purposes. Sometimes we use *preliminary* assessment. Sometimes, at the beginning of the school year we are ready to move ahead with our curriculum, but are uncertain as to how well the students are prepared. Let's say that a 5th-grade teacher has learned from previous experiences that she cannot always begin the school year with the prescribed 5th-grade math curriculum. She has found that sometimes her students came from classes in the previous year where the 4th-grade teacher did not complete the prescribed materials or that some students forgot a great deal over the summer. Therefore, within the first two weeks of the school year, our 5th-grade teacher gives her students a test of the prerequisite math skills, the skills that they need to be prepared to master the new skills that she plans to teach them. If her students perform well on this preliminary test then she knows that she can safely move ahead with the prescribed lessons. However, the test is likely to reveal that there are a few skills that she will need to teach first, or at least review, before she can begin the prescribed curriculum.

A variation of this is the *placement* assessment. Placement assessments are very commonly used in colleges where incoming freshmen are often given placement assessments in a variety of subjects. For example, many colleges require students to take one or more math courses and frequently require proficiency with basic algebra. Therefore, they give mathematics placement tests to their incoming freshmen. If students' scores are high

enough, they are frequently exempt from taking the freshman algebra class and can take a higher-level mathematics class. These placement tests are especially important in fields like science and engineering where students are typically expected to take a calculus sequence. Are students ready to take Calculus I or do they need to take the Pre-Calculus course first? Placement tests are also used in other areas such as English and foreign languages, where sequenced courses are often available. Placement tests are also used in elementary schools, middle schools, and high schools. Sometimes they are used at the end of the school year to help determine next year's placements.

Definition

Preliminary or **placement** assessments are assessments performed within the first two weeks of the semester that are designed to measure students' prerequisite skills.

Diagnostic Assessment

A second type of assessment is the *diagnostic* assessment. We use diagnostic assessments when we recognize that students are not performing well, but are uncertain of the exact problem. For example, Ms. Kindya, a 2nd-grade teacher, notices that Tim is struggling with reading, yet she is unable to identify clearly which skills are giving Tim the greatest problems. Luckily, Ms. Kindya is familiar with and skilled at using a variety of reading diagnostic tests. She chooses the test that appears most appropriate, administers it, and is able to identify the specific skills that are giving Tim trouble. She now can provide Tim with the specific help that he needs. We are likely to use diagnostic tests most frequently in the areas of elementary reading and mathematics. However, we also have diagnostic tests in a number of other fields.

Although some diagnostic tests are designed for classroom teachers to administer, others require considerable training. In those cases, you may need to call on a specialist, like a reading specialist, a speech-language pathologist, or a school psychologist, to administer such tests.

Definition

A **diagnostic** assessment is any type of assessment used to identify an individual student's learning decrements.

Formative Assessment

The third type of assessment that we use is known as *formative* assessment. Formative assessment is any type of assessment device that we use while an instructional unit is in progress. It is used primarily to give the teacher feedback on how the unit is progressing. Let's say that you are a 5th-grade teacher and are teaching a six-week unit on adding and subtracting fractions. You might recall that in order to add or subtract fractions, you must first find a common denominator. For example, to complete this problem $\frac{1}{2} + \frac{1}{4}$ you must convert the $\frac{1}{2}$ to $\frac{2}{4}$. Therefore, the first two to three weeks of the unit are

dedicated to teaching the students this prerequisite skill, finding the common denominator. Before you move into the part of the unit where the students actually learn to add and subtract fractions, you want to make certain that they are doing very well with this prerequisite skill. Therefore, at the end of week two you give your students a pop quiz that requires them to complete 10 common denominator problems. Over the weekend you score the quizzes. In this case, you are using the quizzes primarily as a feedback device for yourself. You want to find out if the class is showing sufficient proficiency with this skill to move to the next step? If they performed well enough on the quiz, then you can move to the next step early next week. If, however, they are still not performing well with the task, it means that you need to spend several more days working with common denominators before you can move on. Formative assessments are designed to give you, the teacher, feedback. Therefore, most psychometricians argue that formative assessments should not be used to give students grades. Others, however, claim that unless assessments are graded students may not take assignments seriously.

Besides pop quizzes, there are many other techniques that can be used for formative assessment. If we continue with the math instruction example, some other techniques that could be used would include having the students perform seat work while you walk around the room looking at their work. You could also have them come up to the board and solve problems there. You could play a game where they have to find the common denominator. There are a large variety of techniques that you can use to gauge students' progress. Many formative assessment techniques involve informal assessment.

Definition

Formative assessment is any type of assessment device that we use while an instructional unit is in progress and is used primarily to give the teacher feedback on how the unit is progressing.

Summative Assessment

The fourth type of assessment is *summative* assessment. Summative assessments are performed at the end of chapters or units to determine the students' level of competency with the material and to assign grades. Many of you may have assumed that summative assessment was the only type of assessment used in school. It is the type that we probably recall most often and with which we have had the most experience. Not all teachers use preliminary or diagnostic assessment. Even when our teachers were using formative assessment we, as students, were not always aware of why they were doing it. However, we were typically keenly aware of summative assessment.

Definition

Summative assessments are performed at the end of chapters or units to determine the students' level of competency with the material and to assign grades.

All four types of assessment play important roles in the classroom. However, in many

classrooms formative assessment is the most common type of assessment, as it should be. Elementary school teachers tend to use formative assessment often, especially with the youngest students. As students move into the higher grades, however, there is a tendency to use it less often, which may be a mistake. Many college instructors hardly use it at all. At times, after working with my class on a difficult topic, I find myself asking them this question, “Is that clear to everyone?” As they nod their heads, I feel satisfied that they understand the material and move on to the next topic. I often think that I’ve completed an effective formative assessment. In the last few years, I have come to recognize that when I do that, I am using a very ineffective approach. How many times, in your experiences, has someone asked you a question like that after teaching you something fairly complex? You may have responded in the affirmative, only to discover later, when you tried to apply the new learning, that you really did not understand. It is much more effective to actually have students try out the new learning to assess if they can use it correctly.

MAXIMUM vs. TYPICAL PERFORMANCE MEASURES

When we measure some skills and abilities we want to assess how well our students can perform. We want to assess them at their best. However, when measuring other types of traits, we are more interested in students’ typical behaviors. Classroom tests designed to measure a student’s level of competence with some skill should be geared toward measuring the student’s maximum performance. We want to assess them doing as well as they can do. However, personality tests or career interest inventories and similar measures work best when they measure a student’s typical behavior or interest.

There are differences in how we design the most effective maximum performance measures compared to how we design effective typical performance measures. If this text were designed to be used in a course on psychological measurement, we would spell out those differences. However, since this text is designed primarily for classroom teachers, that is not really necessary. The tests and other measures that classroom teachers use will, almost exclusively, be maximum performance measures. That will be our focus in this book.

Definitions

Maximum performance measures are used when we want to assess how well our students can perform.

Typical performance measures are designed to assess a student’s most common, or typical, behavior.

USES OF MEASUREMENT

Since this chapter is introducing the section in this part on measurement, this might be the appropriate place to let you know what to expect in the remaining chapters. Chapter 2, Frames of Reference, will introduce you to the various ways in which we can interpret test scores. Chapter 3, Developing Objectives, covers the various approaches

that educators have used to establish goals and objects. We use objectives to help you decide what you will teach and what you will test. Chapter 4, Reliability, will introduce you to the concept of reliability, consistency within tests. What is reliability? How do we measure it? How can we develop tests that are reliable? Chapter 5, Validity, deals with the issue of whether a test is actually measuring what we think it should be measuring.

Chapters 6 through 11 deal with building effective classroom tests. Chapter 6 discusses short-answer and completion items. Chapter 7 deals with essay items. Chapter 8 deals with multiple-choice items. Chapter 9 deals with true–false items and other alternative choice variations. Finally, Chapter 10 discusses how to assemble the individual items into a test. Chapter 11 is a follow-up on the earlier chapters on classroom testing. It deals with how to analyze your tests to assess if they are performing as you had hoped they would.

Chapters 12, 13, and 14 deal with alternative forms of assessment. Chapter 12 delves into informal assessment, which was briefly introduced in this chapter. Chapter 13 deals with performance assessments, which we use when we need to rate a student’s actual performance as a judge would do when rating an athletic competition such as gymnastics. Chapter 14 discusses how to use portfolios effectively.

The last three chapters deal with other important topics. Chapter 15 deals with teaching students to be effective test takers. Chapter 16 describes the various types of standardized test you are likely to encounter in schools. Finally, Chapter 17 deals with many alternative ways to report test scores and why many of these are, at times, better than reporting raw scores.

SUMMARY

In this chapter we examined the various roles that measurement plays in the classroom. We also differentiated among three commonly used terms: assessment, measurement, and evaluation. We differentiated between formal assessment and informal assessment. Classroom tests, quizzes, and similar devices are frequently referred to as formal assessment, whereas teacher questions, observations, and many other techniques are often referred to as informal assessment. We differentiated among four ways that we use assessment. We use preliminary or placement assessment to identify students’ prerequisite skills. We use diagnostic assessment when students are experiencing learning problems. We use formative assessment to judge how well our instruction is proceeding. We use summative assessment to judge our students’ competence with the material that they have been taught. Finally, we differentiated between maximum and typical performance measures. In the classroom we typically use maximum performance measures because we want to assess how well our students can perform the new skills that they have learned. In other settings, especially when examining concepts like personality, we use typical performance measures which will identify the individual’s typical, or usual, behavior.

EXERCISES

Identify the type of assessment being used.

Formal or Informal?

1. A 5th-grade standardized reading test.
2. Verbally asking a group of 3rd graders comprehension questions about a story they just read.
3. A pop quiz in a 7th-grade math class.
4. Observing the students as they work in small groups.
5. Using a checklist while a 4th-grade class gives oral presentations.

Preliminary or Placement?

1. A 3rd-grade teacher asks her students multiplication facts before starting a unit on multiplication to assess if they have any experience in working with that operation.
2. In a 2nd-grade classroom, the students take a reading test to find their reading level so the teacher can accurately place the children in ability groups.
3. College freshmen at the University of Oklahoma are required to take a writing test. The test is evaluated by a group of professors and the students are then put in the appropriate composition classes.

Diagnostic, Formative, or Summative?

1. At the end of a unit on weather, Mrs. Jones gives her 4th-grade class a test to assess how much they have learned. This test is 20% of their semester grade.
2. Mr. Gregory is a 1st-grade teacher. He is introducing his children to bar graphs. He is not sure that all of the students understand the material. He has the students work in groups to create a simple bar graph. He observes the children as they work and collects the graphs to check for accuracy.
3. Joey is struggling in all subject areas. He is a bright child and therefore his teacher, Mrs. Buckwalter, believes it is a reading disability. She gives him multiple reading evaluations to assess where the problem lies.
4. Each week Mr. Grables' students receive 15 spelling words. At the end of the week they take a test on the words and their scores are recorded.
5. Mrs. Bacco is teaching a lesson on social studies. She plays a review game a week before the unit test. She observes the students as they answer questions and makes notes on the concepts they will need to review before the test.

SPOTLIGHT ON THE CLASSROOM

Miss Anderson is a first-year 3rd-grade teacher at Wilkins Primary. She is beginning to teach a unit on measurement. Measurement is not directly taught to students in her school before 3rd grade but she knows that some students have picked a few things up on their own. Before she begins, Miss Anderson gives the students a worksheet to complete. The worksheet questions their basic knowledge of measurement. She uses the information to plan her lessons based upon what the students already know.

Miss Anderson decided to introduce a new unit of measurement—such as temperature, length, and weight—one at a time every few days. At the end of each lesson she gives the students a small quiz before moving on to another unit of measurement.

At the end of the unit, Miss Anderson plays a measurement review game with her students. She asks the students questions and they work in groups to come up with answers. She uses this game the day before their unit test. Miss Anderson’s unit test covers all of the units of measurement she has taught. She records the students’ scores on the test. The test is worth 100 points. A score of 90–100 is an A, 80–89 is a B, 70–79 is a C, 60–69 is a D, and below 60 is considered failing. The scores are 10% of the students’ final math grade for the semester.

1. What forms of assessment were used in the above scenario?
2. What changes, if any, would you make if you were teaching this unit?

STUDY TIPS

Setting Effective Academic Goals

Researchers such as Locke and Latham (2002) have demonstrated that one of the most effective ways for students to stay motivated and complete tasks is for them to set effective academic goals for themselves. Perhaps, at the beginning of a term, you may have said to yourself that you wanted to achieve a 3.25 grade point average (GPA) this term or that you wanted to earn an A or a B in a particular course. If you have ever said something like that you were on your way to developing a goal. However, unless you developed a plan concerning how you would achieve that end, you only took the first step in setting an effective academic goal. I have even heard it suggested that such a statement is merely a wish and not a real goal. Goals require planning.

Whenever you begin any activity, it often helps if you spend some time setting goals. Goals can serve many different purposes, but here you will be using goals to help you maintain your motivation. Locke and Latham (2002) point out that when it comes to effective academic goals, the best goals have three characteristics. They are specific, proximal, and moderately challenging.

At the beginning of this term you may have said to yourself, “I would like to do better this term than I did last term.” Although there is nothing intrinsically wrong with that goal, it is quite general. If you earned a 2.25 GPA last term, would you be satisfied with a 2.26 GPA this term? After all, you would have met your goal! However, it is more likely that you really wanted to see a more substantial improvement in your academic record. Therefore, it would be even more effective if you set a more specific goal. Perhaps you could set a goal to earn a 2.75 GPA this term. You could then determine what grade you would need to earn in each course to achieve that goal. A specific goal is more motivating than a very general goal.

You set some goals for the distant future (distal goals). Perhaps you decided that you want to be a millionaire by the time that you are 50. There is nothing wrong with that goal. However, if you are currently 22, then you could easily decide that you don’t even have to start worrying about that goal until you are closer to 40 and, therefore, it is not an effective motivator for you at the present time. You can set other goals for the near future (proximal goals). For example, you may decide to finish reading this chapter by 7:00 p.m. this evening. When you set proximal goals (also sometimes known as “near”

goals) they can be very motivating. Although there are other tasks that you could perform this evening, it is very likely that you will complete this task first.

A third characteristic of effective motivational goals is that they are moderately difficult or moderately challenging. If you set goals that are not challenging, even when you achieve them, you get no sense of satisfaction. On the other hand, if you set goals that are extremely difficult or challenging, or goals that are simply unrealistic, then you may fail to achieve those goals and will feel like a failure. The best goals are moderately challenging, but achievable. With those types of goal, you have a good chance of being successful and will have a sense of accomplishment once the goals are completed.

Having goals that are specific, proximal, and moderately challenging is a good start. However, it is not enough! You then need to plan the steps that you will need to take to achieve those goals. For example, if you set a goal to earn an A in this course, you will need to develop a plan to make that goal achievable. How much study and preparation time do you need to allocate to this course on a weekly basis in order to achieve that A? Do you have all the prerequisite skills required to earn an A or will you need to arrange for some assistance? Should you arrange for a tutor, join a study group, or do some additional preparation for this course?

With effective goal setting and planning you are much more likely to stay motivated and are also much more likely to reach your academic goals.¹

NOTE

1. This material has been adapted from Van Blerkom (2009).