14

Error

The Scourge of Research

Recall the assumptions required to interpret regression coefficients (paths) as estimates of effects of one variable on another:

1. There is no reverse causation; that is, the model is recursive.
2. The exogenous variables are perfectly measured, that is, they are completely reliable and valid.
3. A state of equilibrium has been reached. This assumption means that the causal process has had a chance to work.
4. No common cause of the presumed cause and the presumed effect has been neglected; the model includes all such common causes (Kenny, 1979, p. 51).

We have dealt with several of these assumptions, such as the effect of neglecting a common cause, and I promised we would return to assumption 2: the assumption of perfect or near perfect measurement of the exogenous variables. Obviously, this assumption is violated routinely—perfect measurement is rare to impossible—but what effect does this violation have on our research? In addition, inaccurate measurement of the endogenous variables also affects estimates in path models.

It is worth noting that issues of reliability and validity of measurements affect *all* research, not just that based on path analysis and multiple regression. Many of us think of measurement as separate from statistics, but they are inexorably intertwined. In a laboratory experiment our experimental conditions (the exogenous variable) may be clear-cut and thus perfectly

measured (e.g., treatment versus control), but the dependent (endogenous variable) (e.g., a measure of self-esteem) may be considerably less reliable. This lack of reliability may result in an underestimation of the effect of the experimental treatment, with even a truly meaningful finding showing up as statistically nonsignificant. In applied research, there may be variations in the treatments by those responsible for providing the experimental treatment. Teachers in an experiment designed to compare the effects of two methods for teaching reading may use other methods outside the experimental procedure. This variation is, in fact, unreliability and invalidity in the independent (exogenous) variable, which will also cloud the results of the research. In fact, the effect of measurement on decision making affects *every* aspect of life. Your physician may prescribe or not prescribe medication for high blood pressure depending on her measurement of your blood pressure; if her measurements are unreliable, however, you may receive unnecessary treatment or not receive needed treatment. You may have costly repairs completed on your car based on unreliable measurement, and so on. Measurement accuracy affects all research and all decisions made from these measurements. Why, you may wonder, does it?

EFFECTS OF UNRELIABILITY

The Importance of Reliability

In classic measurement theory, we might administer a test, or survey, or other measurement to a group of people. There will be variation in their scores; some people will score high, some low. We also know that there will be error in their scores; all measurement involves error. This aspect of scores is represented in Figure 14.1. V represents the total variation in a set of scores on some measurement. This total variance can be divided into variation due to error (V_e) and true score variation (V_t): $V = V_t + V_e$. Using this definition, reliability is the proportion of the true score variance to the total variance: $\frac{V_t}{V}$. This makes sense: the greater the error in a set of scores, the less a person's score on that measure is a result of true variation and the less reliable the measurement.

Figure 14.2 illustrates the effects of unreliability in path analytic format. In this graphic, a person's score on any measurement is affected both by the person's *true score* and by *errors* of

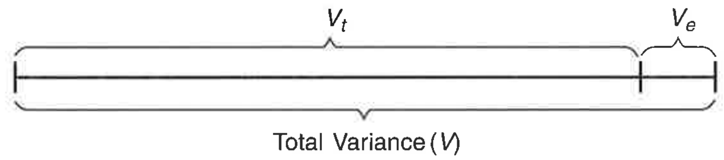

Figure 14.1 Variance definition of reliability. Reliability is the proportion of true score variance to total variance ($\frac{V_t}{V}$).

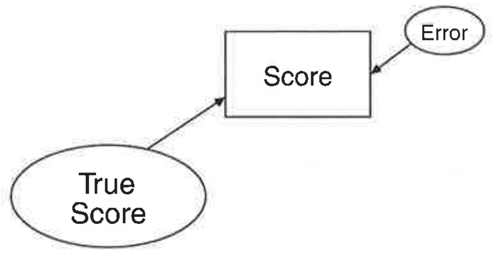

Figure 14.2 Path analytic definition of reliability; a person's score on a test or measurement is affected by their true, but unknown, score and by error.

measurement. In this graphic, error is equivalent to V_e and the true score to V_t. Note that the actual, measured score is the only measured variable in this model; both the true score and the error are unmeasured and unknown.

The reliability of a test, scale, survey, or other measure places an upper limit on the correlation that the measurement can have with any other measurement. As a general rule, a second variable will correlate with the measured score through correlation with the true score. That is, other variables will generally correlate with the V_t portion of the variable illustrated in Figure 14.1, not the V_e portion. This, then, is the reason that measurement quality affects statistics and research: a less reliable measurement limits the correlations a variable can have with any other variable. Since correlations are the statistic underlying multiple regression, path analysis, ANOVA, and other derivatives of the general linear model, unreliable measurement causes us to underestimate the effects of one variable on another in *all* these methodologies.

Effects of Unreliability on Path Results

What effect does measurement error have on path analytic results? Figure 14.3 shows the results for the homework model from Chapter 13. In this model, whether we realize it or not, we are assuming that all the variables in the model are measured without error, with perfect reliability. As researchers, we may recognize that the variables in the model are measured with different degrees of error, but the model assumes they are all error free.

Focus on the variable of homework. Homework is based on student self-report of the average amount of time students spend on homework in several academic areas. Undoubtedly, error is inherent in this variable, not only because of the self-report nature of the questions, but also because, perhaps more importantly, students were asked to approximate their average amount of time per week. I would not be surprised to discover that this variable had a reliability of only about .70, with a corresponding error of 30%. If we build such estimates into the path model, what will be the effect on the estimates of paths?

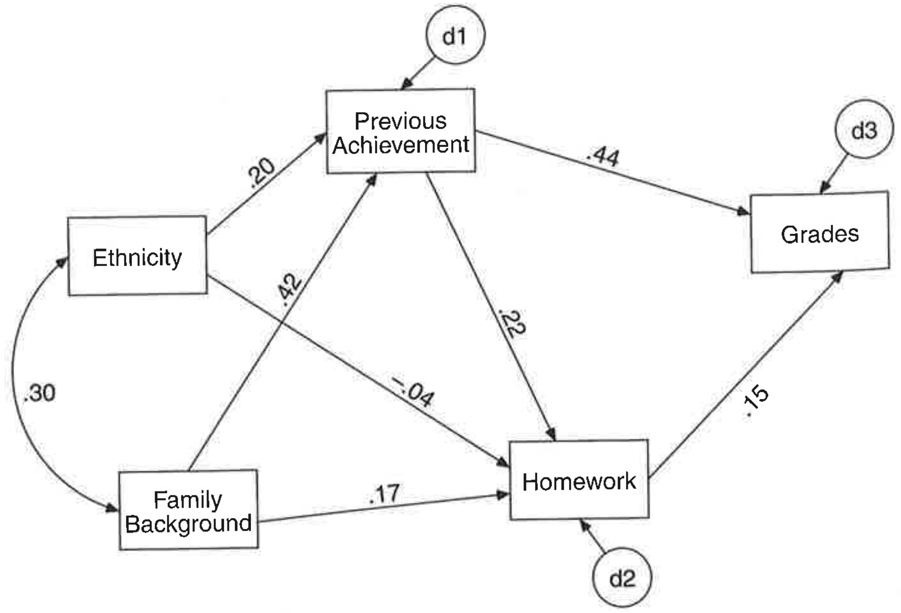

Figure 14.3 Homework model from Chapter 13 revisited.

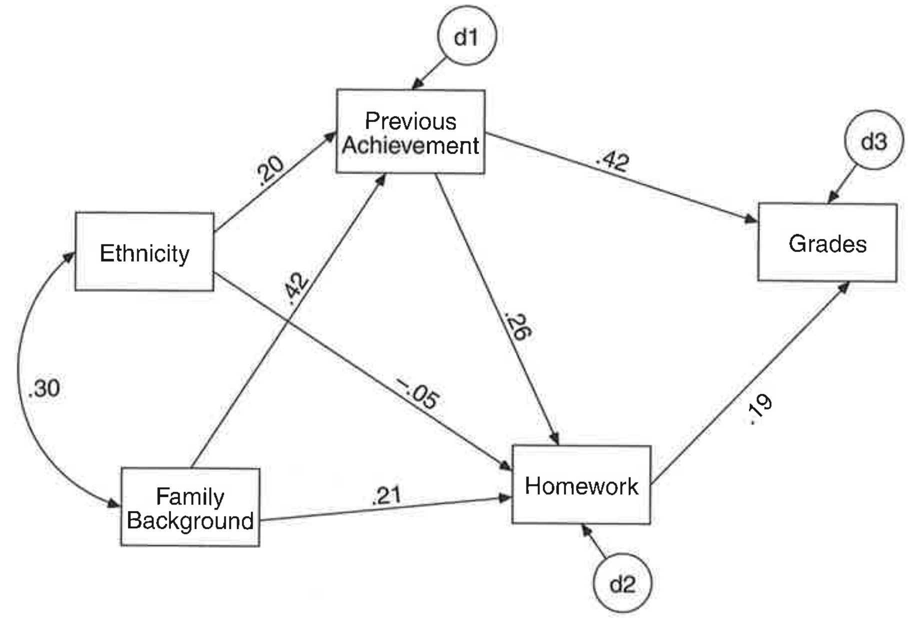

Figure 14.4 Effects of error. This model recognizes and accounts for the unreliability in the Homework variable; with this recognition, the apparent effect of Homework on Grades increases.

Figure 14.4 shows a model that recognizes this unreliability (reliability = .70, error = .30) in the Homework variable. Note the increase in the apparent effect of Homework on Grades, from .15 in Figure 14.3 to .19 in Figure 14.4. What this means is that when we assumed that the error-laden Homework variable was perfectly reliable, as in Figure 14.3, we *underestimated* the true effect of Homework on Grades. In contrast, when we recognize the error inherent in this variable, we obtain a more realistic and larger estimate of the effect. This is also the most common effect of error in models: unreliability artificially reduces our estimates of the effects of one variable on another.

Note also that many of the other paths in the model are different from those in Figure 14.3. Indeed, all paths to Homework increased in magnitude, and the path from Achievement to Grades decreased slightly. Recognition of the error that exists in the Homework variable resulted in changes in many of the paths in the model.

But Homework is not the only less than perfectly reliable variable in the model. What about Grades? Grades were also based on student self-report, plus there are well-known problems with Grades as measures of student learning, including variations in grading standards from teacher to teacher, the unreliability of teacher-made tests and other components of grades, and the likely clouding of other variables (e.g., students' apparent interest) in teachers' grading practices. Given these deficiencies of Grades, it is probably reasonable to estimate their reliability at a maximum of .80 (and 20% of the variation in scores due to error).

Figure 14.5 shows the results of recognition of this level of error for the Grades variable (assuming perfect reliability for the other variables in the model). In this model, compared to Figure 14.3, the magnitude of the paths to Grades from both Previous Achievement (from .44 to .50) and Homework (from .15 to .17) increased.

Although it is not obvious from these figures, the effects of unreliability are different depending on whether the variable in question is exogenous or endogenous. Briefly, error

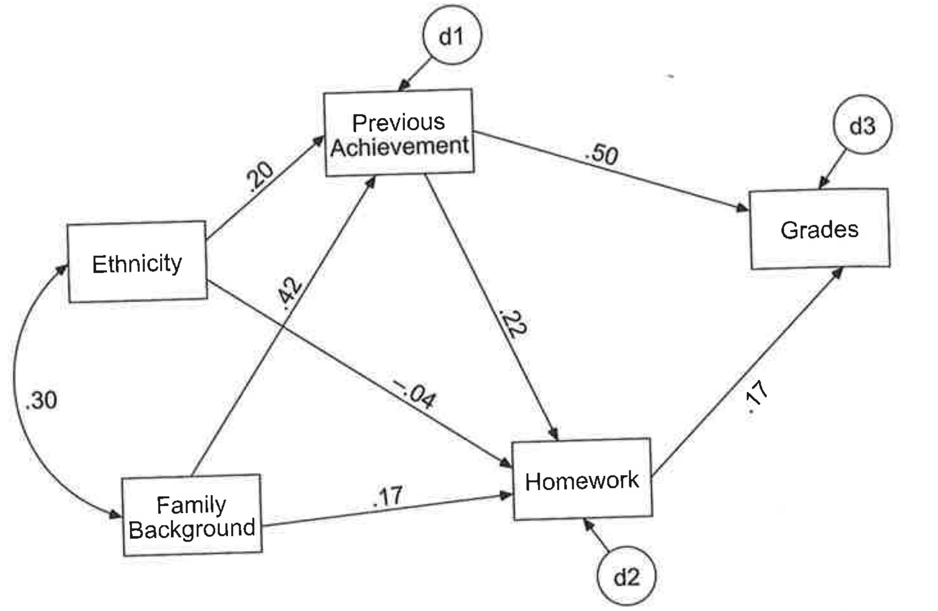

Figure 14.5 Effects of error. This model shows the result of recognizing the error inherent in the Grades variable.

in an exogenous variable affects both the standardized and unstandardized paths, as well as their statistical significance. Paths from other exogenous variables (in addition to the error laden one) may be affected. In contrast, error in an endogenous variable affects only *standardized* estimates of effects, leaving unstandardized effects unchanged. The unstandardized paths for the model shown in Figure 14.5 would be the same as those for the model shown in 14.3, despite the differences in the standardized paths. This difference is why error in exogenous variables is more consequential than error in endogenous variables. When a variable is in the middle of a model—exogenous in relation to some variables, endogenous for others—the results of error are more complex, as in the example recognizing error in Homework (Figure 14.4). The bottom line is that measurement error affects estimates of effects, but is more serious for exogenous variables [for more information, see Bollen, 1989 (chap. 5); Rigdon, 1994; or Wolfle, 1979].

These examples have corrected for unreliability in a single variable. What would happen if we were to recognize the unreliability in *all* the variables in the model? If you think about it, all the variables in the model are unreliable to one degree or another. Even Ethnic orientation, probably the most reliable variable, likely has some error. Students may not read the survey question accurately, students who could legitimately claim to belong to more than one ethnic group are allowed only one answer, some students simply knowingly mark the wrong response, and there may be errors in coding of students' responses. For whatever the reason, even this variable likely includes some error.[1]

The model shown in Figure 14.6 attempts to recognize the error inherent in every variable in the model. For this example, I assumed that error was responsible for 30% of the variability for Homework, 20% for Grades, 5% for Ethnicity, 20% for Family Background, and 10% for Previous Achievement. These are plausible estimates. Note that every parameter estimate in the model changed from those shown in Figure 14.3. Most estimates increased in magnitude, but one, the path from Ethnicity to Previous Achievement, decreased (from

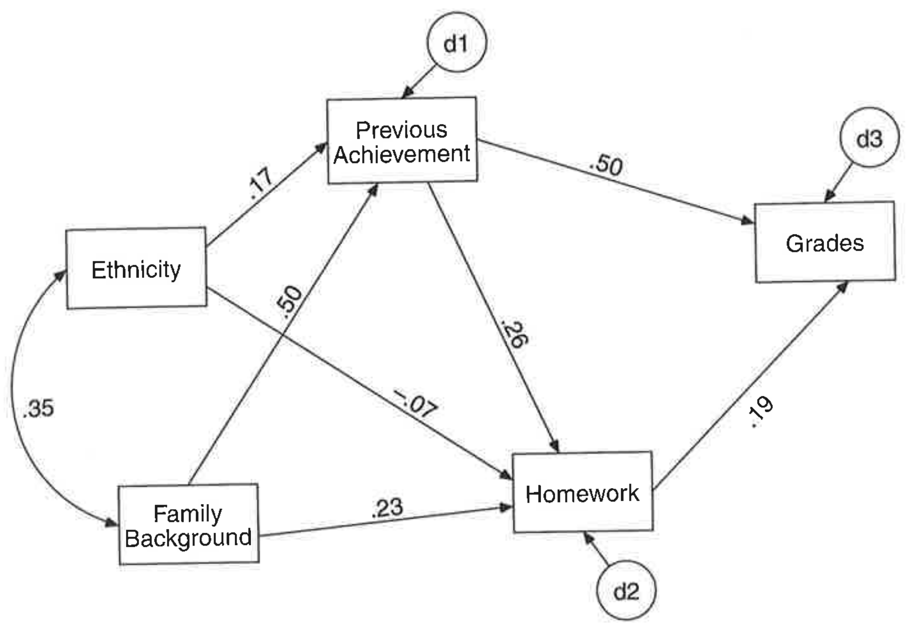

Figure 14.6 Effects of error. This model recognizes the error inherent in all the variables in the model. Compare the coefficients here with those shown in Figure 14.3.

.20 in Figure 14.3 to .17 in Figure 14.6). Recognition of the error inherent in the variables in our models will often, although certainly not always, result in larger estimates of the effects of one variable on another. With such complex patterns of errors, estimates may increase, decrease, or stay the same.

These examples illustrate the effects of measurement error on estimates of the influence of one variable on another in path analysis (as well at MR, ANOVA, etc.). What can researchers do to avoid misestimating such effects? We can strive for better measures, but no measures are error free. We could also correct the correlations for all the variables in the model using estimates of each variable's reliability and the common formula for correcting for attenuation, $r_{T_1T_2} = r_{12} / \sqrt{r_{11} \times r_{22}}$, where $r_{T_1T_2}$ is the corrected, or "true" correlation, r_{12} is the original correlation, and r_{11} and r_{22} are the reliabilities of the two variables. This solution is not very satisfying for several reasons. First, it divorces the correction from model testing; indeed, the process smacks of statistical voodoo. Second, when there are multiple estimates of reliability, such as with several studies providing estimates, it is unclear which estimate should be used. Conversely, no estimates of reliability may be available for a given measure. Finally, although this method might deal with unreliability of measures, it ignores problems of invalidity.

EFFECTS OF INVALIDITY

The Meaning and Importance of Validity

What effect does invalidity have on estimates of effects? In classic measurement theory, validity may be considered as a subset of reliability. An example will illustrate how these measurement concepts are related. Suppose that you are interested in the effects of reading comprehension on subsequent delinquent behavior. One task is to measure reading comprehension. You will find that different tests of reading comprehension use different methods of measurement. Test 1, for example, asks research participants to read a passage on one page

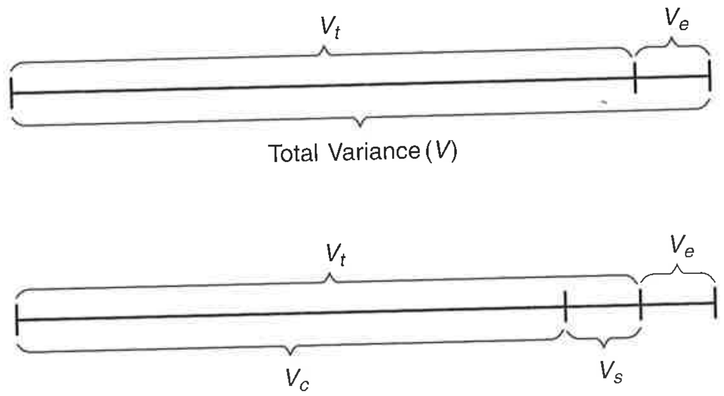

Figure 14.7 True score variance may be further subdivided into common variance (V_c) and specific or unique variance (V_s). Validity is related to common variance.

and then point to one picture (out of four choices) on the next page that best illustrates what they read in the passage. Test 2, in contrast, asks participants to read a passage (e.g., "stand up, walk around the table, then sit down") and then do what the passage requested. Test 3 uses a "cloze" procedure; the participant reads a passage with one or several words missing and then supplies the missing words based on the meaning of the text.

It is clear that each of these tests measures reading comprehension to some degree. But each test also measures something else, something other than reading comprehension. Test 1 also measures the ability to translate something read into a picture; Test 2 measures the ability to act out something read; Test 3 measures the ability to pick from one's knowledge store the word or words that will make the most sense when inserted in a passage. Each test may measure these unique skills reliably, but these skills are not the same as reading comprehension.

We are also not interested in the variation in scores due to these unique skills. We are interested in the effects of *reading comprehension* on delinquent behavior, not the effects of the ability to translate text into mental pictures (Test 1) or the unique skills measured by other tests on delinquent behavior. This variation due to these unique skills will not be removed through correction for attenuation, however, because these skills are measured reliably and are not due to error.

As shown in Figure 14.7, it is possible to extend the earlier variance definition of reliability. The true score variation (reliability) can be divided further. Using the reading comprehension example, one component of the true score variation for each test is the variance that these three tests have in common, the common variance, or V_c. What do the three Reading Comprehension Tests measure in common: reading comprehension! Each test also measures something unique or specific, however, and this component of the true score reliability is symbolized as V_s, for specific variance. The common variance, V_c, is an estimate of the validity of each test and thus demonstrates that validity is a subset of reliability. The V_s, the unique or specific variance of each test, is sometimes called the *specificity*, or the unique variance. For our present purposes, it represents invalidity and needs to be taken into account in our research on the effects of reading comprehension on delinquent behavior.

Accounting for Invalidity

How can we take this invalidity into account? Another way of conceptualizing the problem is as a path model, as shown in Figure 14.8. The diagram illustrates the influences on individuals' scores on the three Reading Comprehension Tests. Each person's score on each test is first

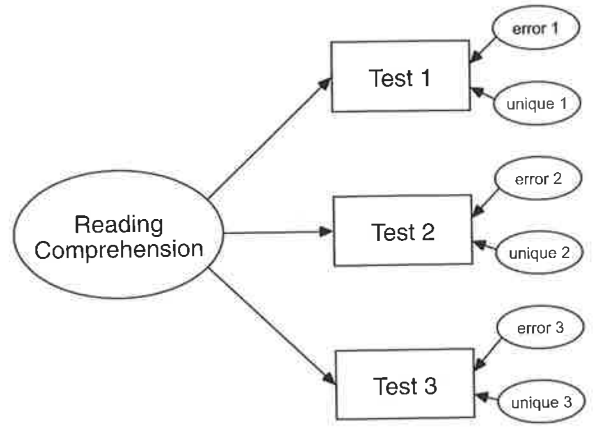

Figure 14.8 Using path models to understand validity. Individuals' scores on three tests of Reading Comprehension are affected by their true level of Reading Comprehension and by error and the unique aspects of each test.

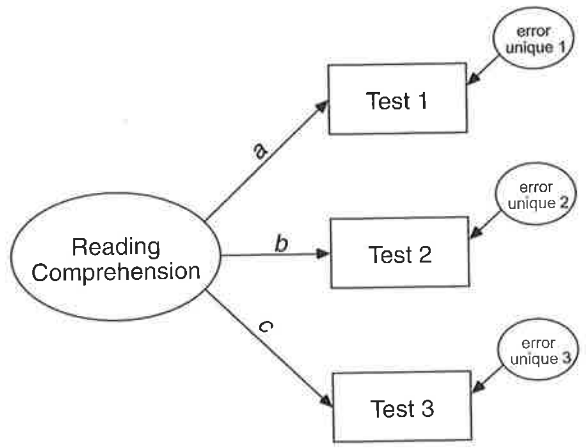

Figure 14.9 Reading Comprehension measurement model; we can generate equations to solve for the paths from Reading Comprehension to the three Tests.

affected by his or her level of reading comprehension. Reading Comprehension—the true level of reading comprehension—is an unmeasured or latent variable and is thus enclosed in an oval. Each person's scores on each test are also affected by error (unreliability) and by that person's level of the unique skills measured by each test (one's ability to translate text into pictures, and so on). These are also unmeasured variables. Our primary interest, of course, is in the Reading Comprehension latent variable.

Figure 14.8 is just another path model, and we can solve it in much the same way we solved the path models in Chapter 11. Figure 14.9 shows a slight revision of the model, with the error and unique variances combined for each variable and the paths labeled to help develop equations. Figure 14.10 shows the correlations among the three tests. As in Chapter 11, we can use the tracing rule to develop equations:

$r_{12} = ab$,
$r_{13} = ac$, and
$r_{23} = bc$.

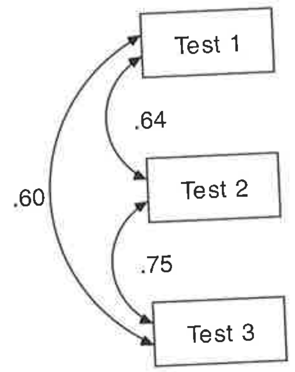

Figure 14.10 Correlations among the three Tests used to solve for the paths.

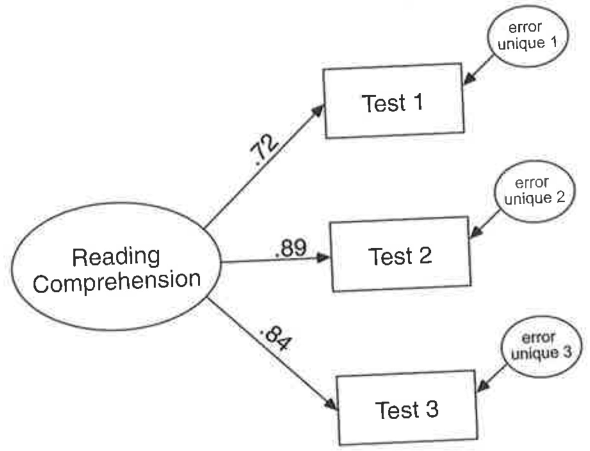

Figure 14.11 Solved Reading Comprehension measurement model.

If we combine the first two equations, we get $r_{12}r_{13}=abac$, which can be simplified as $a^2bc = r_{12}r_{13}$, or $a^2 = r_{12}r_{13} /bc$. Because $bc = r_{23}$ from the third equation, $a^2 = r_{12}r_{13} /r_{23}$ and $a=\sqrt{r_{12}r_{13} / r_{23}}$ We can also solve for b and c: $b=\sqrt{r_{12}r_{23} / r_{13}}$ and $c=\sqrt{r_{13}r_{23} / r_{12}}$. If you substitute the correlations in these equations, $a = .716$, $b = .894$, and $c = .839$. Figure 14.11 shows the model with the path estimates inserted.

Interestingly, what we have done by solving for the paths in Figure 14.11 is a simple (confirmatory) factor analysis. Figure 14.12 shows output from a factor analysis of these three items in SPSS; the factor loadings from the output are the same as the paths from the Reading Comprehension latent variable to the three reading Tests.[2] The example nicely illustrates the thinking underlying factor analysis: there is a latent, or unmeasured, variable, or factor, that affects individuals' scores on these three Tests and does so to different degrees. The example also illustrates the equivalence of several terms. What we have been referring to as latent or unmeasured variables are equivalent to the *factors* from factor analysis. These latent variables or factors are also much closer to the *constructs* we are interested in than are our normal, error-laden measurements.

Our primary interest, of course, was the influence of Reading Comprehension on Delinquent Behavior. Because we can solve the model to estimate the Reading Comprehension latent variable, we could also use the latent variable in an analysis of the effects of Reading

Factor Matrix[a]

	Factor
	1
TEST_1	.716
TEST_2	.893
TEST_3	.839

Extraction Method: Principal Axis Factoring.
a. 1 factors extracted. 11 iterations required.

Figure 14.12 Reading Comprehension measurement model solved via factor analysis. Our measurement model is a (confirmatory) factor analysis.

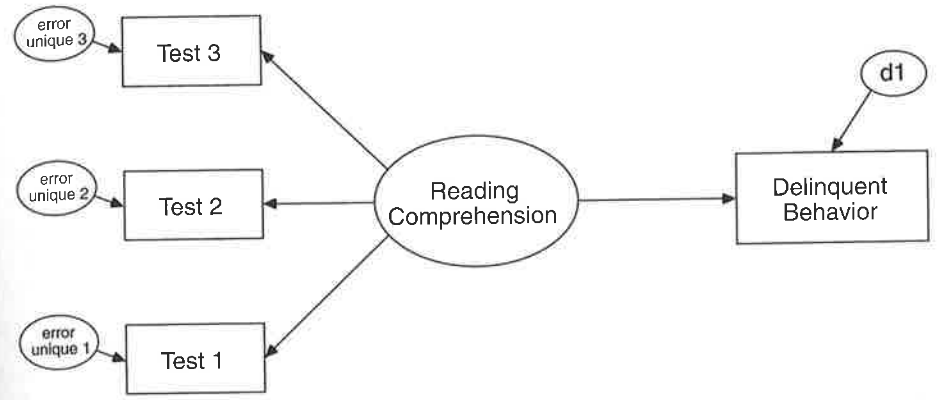

Figure 14.13 We could use the Reading Comprehension factor, or latent variable, in a structural equation to more accurately determine the effect of Reading Comprehension on Delinquent Behavior.

Comprehension on Delinquent Behavior, as in Figure 14.13 (once we were able to measure Delinquent Behavior).

LATENT VARIABLE SEM AND ERRORS OF MEASUREMENT

To return to our more general problem, perhaps this means that the solution to the problem of less-than-perfect measurement is not to correct all the correlations for attenuation but to obtain multiple measures of each construct in our path model, separately factor analyze these items, and then use the factor scores in our path analyses, rather than the original items or tests. This process will rid our measures of both invalidity and unreliability (because ridding the measure of invalidity will rid it of unreliability) and will allow us to get closer to the constructs we are interested in. Although this solution makes sense conceptually, it too has drawbacks. The multistep process separates the different factor analyses (the measurement model) from the testing of the path model (the structural model). It would be preferable to be able to conduct *all* analyses simultaneously.

This is what latent variable SEM does: it performs confirmatory factor analysis and a path analysis of the resulting factors at the same time. In the process, latent variable SEM removes the effects of unreliability and invalidity from the estimation of the effect of one variable on another. By doing so, the method gets closer to constructs we are really interested in. Thus, instead of doing research on the effects of a measure of Reading Comprehension on a measure of Delinquent Behavior, we can come closer to studying the effect of *true* Reading Comprehension on *true* Delinquent Behavior. Alternatively, if we are interested in the effects of income on job satisfaction, we are not interested in the effects of reported income (the number someone reports on a survey) on perceptions of job satisfaction. Instead, we are interested in the effects of *true* income on *true* job satisfaction. In other words, we want to strip away the fog of invalidity and measurement error and get at the true constructs of interest. Likewise, if we are studying the effect of social skills on peer acceptance, we are not really interested in the effects of someone's perceptions of peoples' social skills on their perceptions of acceptance; we are interested in the effects of *real* social skills on *real* acceptance. Latent variable SEM helps us get closer to this level of analysis.

The Latent SEM Model

Figure 14.14 illustrates a generic latent variable structural equation model. To refresh our jargon, latent variables are the same as unmeasured variables or factors. Latent variables are inferred from the measured variables, and they more closely approach the constructs of true interest in the research. Latent variables are enclosed in circles or ovals. Measured variables are also known as observed variables or manifest variables. They are the variables that we actually measure in our research through tests, surveys, observations, interviews, or other methods. Measured variables are enclosed in rectangles. Scores on a reading test, survey items concerning time spent on homework, records of social interactions from playground observations, and a count of errors on a computer task are all examples of measured variables. Actual reading comprehension, time really spent on homework, true social acceptance, and actual mental processing speed are the latent variables we hope to determine through

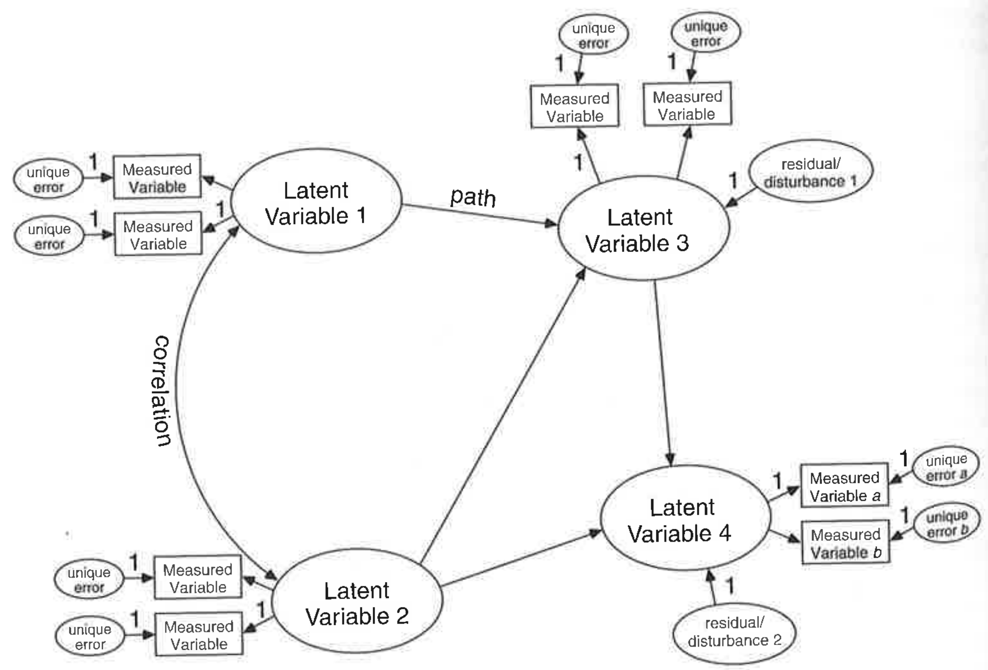

Figure 14.14 Latent variable structural equation model. The model includes a confirmatory factor analysis of the latent and measured variables, as well as a path analysis of the effects of one latent variable on another.

these measured variables. In research we are almost *always* interested in the latent rather than the measured variables, but we often have to settle for the error-laden measured variables as approximations of the latent variables. Not necessarily so with latent variable SEM!

Understanding the Model

The system of paths from the latent to the measured variables is sometimes referred to as the *measurement model*. It is a simultaneous confirmatory factor analysis of all the latent variables in the model. The system of paths and correlations among the latent variables is often referred to as the *structural model*. You can think of it as a path analysis of the latent variables.

You may find it confusing at first glance that both the measured variables and the endogenous latent variables have smaller latent variables pointing to them, but you will soon see that these have previously been defined. Recall that endogenous variables (effects) in a path model have latent variables pointing toward them; these latent variables are generally called either residuals or disturbances. The disturbances represent all *other* influences on the endogenous variables other than those shown in the model. It is the same with *latent* endogenous variables. We need to account for all other influences on the latent variables besides those shown in the model; again we do so with other latent variables known as disturbances or residuals. The small latent variables pointing to the measured variables represent the unique and error variances that we wish to remove from consideration in the SEM as we focus on the true effects of one (latent) variable on another. These unique and error variances are often simply referred to as error or occasionally by Greek letters (e.g., theta delta, theta epsilon), a convention from LISREL. More generally, both types of variables (errors and disturbances) are sometimes referred to as errors.

In fact, you can think of errors and disturbances in the same way. Latent Variable 2 and Latent Variable 3 are not the only influences on Latent Variable 4; there may be a multitude of other such influences outside the model. Residual/Disturbance 2 represents all the other influences on Latent Variable 4 other than those shown in the model. Likewise, Latent Variable 4 is not the only influence on Measured Variable *a*; unique and error variances also affect this and other Measured Variables. "Unique error *a*" represents these influences. Although I will continue to treat disturbances and errors as different, you can thus think of them as "all other influences" on the measured and latent variables.

Figure 14.15 shows a latent variable SEM version of the homework model used in the last few chapters. Note that each variable in the model, except Ethnicity, was measured via multiple measured variables and thus can be estimated by a latent variable. Ethnicity, still indexed by a single item, is still a measured variable in this model. We will explore this model in more detail in subsequent chapters. What is interesting to note at the present time is that the use of latent variables rather than measured variables increased the estimate of the effect of Homework on Grades from .15 (from the path analysis) to above .20 (in the latent variable SEM; the value is not shown in the figure). Again, the latent variable analysis has the advantage of removing measurement error from consideration in the model and thus getting closer to the level of the constructs we are really interested in (e.g., Homework and Grades). The latent variable estimates in this model should thus be the more accurate ones.

We will explore this example in more depth in subsequent chapters. First, however, we will take an important detour in the next chapter into confirmatory factor analysis, or the measurement model portion of latent variable SEM.

Before leaving this chapter, I reiterate that the problems discussed here—the effects of imperfect measurement in research—apply to all research. Here I have focused on the effects of measurement error in nonexperimental research–path analysis and structural equation modeling because this is our focus. But measurement error affects all research, experimental

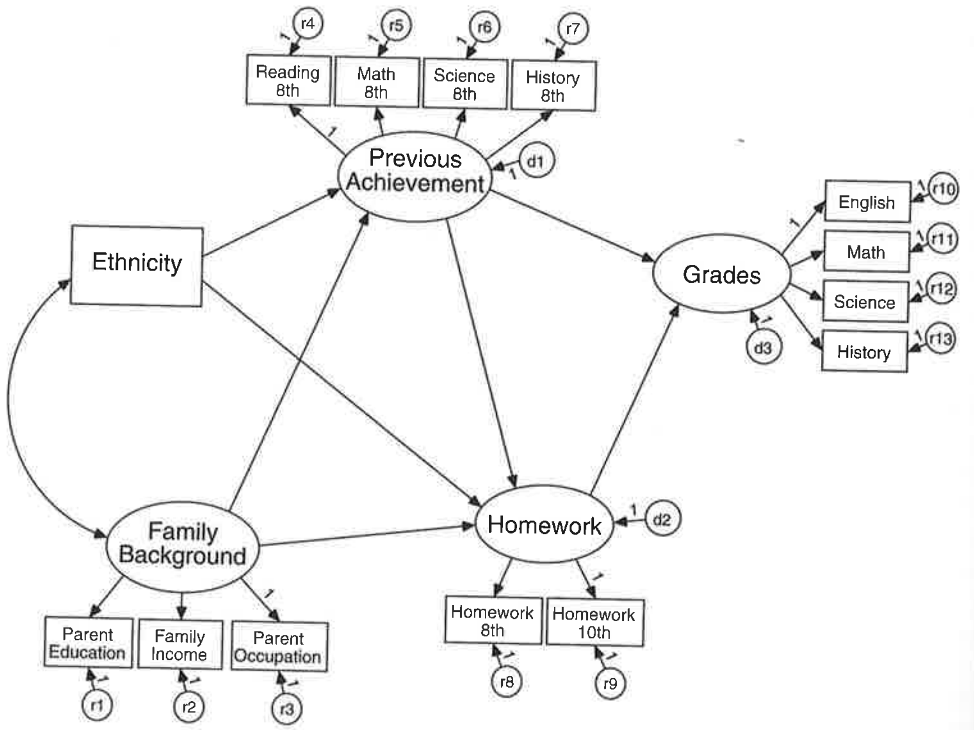

Figure 14.15 A latent variable version of the homework model. All constructs except Ethnicity are indexed by multiple measures. We will examine and test this model in Chapter 17.

and nonexperimental, whether analyzed through ANOVA, correlations, multiple regression, or SEM.

SUMMARY

One assumption required to interpret regression (path) coefficients in a causal fashion is that the exogenous variables be measured without error. We rarely satisfy this assumption and thus need to know the effect of this violation on our estimates of the effects of one variable on another. To expand this discussion, I noted that unreliability and invalidity affect *all* types of research, not just path analysis and multiple regression. Problems in measurement in both the independent and dependent variables affect our research results.

Reliability is the converse of error. Error-laden measurements are unreliable, and reliable measurements contain little error. We can consider reliability from the standpoint of variance by thinking of true score variance as the total variance in a set of scores minus the error variance. In path analytic form, we can think of a person's score on a measurement as being affected by two influences: their true score on the measure and errors of measurement. The true score and error influences are latent variables, whereas the actual score the person earns on the measurement is a measured variable. These concepts are important for research purposes, because other variables generally correlate with the true score, but not the error. For this reason, the reliability of a measurement places an upper limit on the correlation a variable can have with any other variable. Unreliable measurements can make large effects look small and statistically significant effects look nonsignificant.

The path models we have been discussing so far assume that the variables in our models are measured with perfect reliability. In a series of models, I demonstrated what would

happen when we recognized and quantified the unreliability of these measurements. When unreliability was taken into account in these models, the apparent effects of one variable on another changed and usually increased. Taking unreliability into account in our research will improve our estimates of the effects of one variable on another.

Reliability is not the only aspect of measurement that needs to be considered, however; there is also validity. I demonstrated that a measurement may be reliable but may focus on some unique skill, rather than the central skill we are interested in. Said differently, a measurement may be reliable but may not be a valid measure of our construct of interest. As it turns out, validity is a subset of reliability. We can get closer to valid measurement, closer to the constructs of interest in our research, by using multiple measures of constructs.

Latent variable structural equation modeling seeks to move closer to the constructs of interest in our research by using such multiple measures. With latent variable SEM, we simultaneously perform a confirmatory factor analysis of the measured variables in our research to get at the latent variables of true interest, along with a path analysis of the effects of these latent variables on each other. In the process, latent variable SEM removes the effects of unreliability and invalidity from consideration of the effects of one variable on another and avoids the problem of imperfect measurement. In the process, latent variable SEM gets closer to the primary questions of interest: the effect of one construct on another.

Although our discussion focused on the effects of imperfect measurement in multiple regression and path analysis, it is worth remembering that measurement affects every type of research, however that research is analyzed. With the addition of latent variables to SEM, we are able to take measurement problems into account and thus control for them.

EXERCISES

1. Pick a research study in your area of interest. Describe the latent variables, the constructs the authors were interested in. What was the construct of interest underlying the independent variable(s)? What was the construct of interest underlying the dependent variable(s)? What measured variables were used to approximate these constructs?
2. How could you convert this research from a measured variable study into a latent variable study? Think of ways to include multiple measures of the researchers' independent and dependent variables. Draw a model incorporating both measured and latent variables.
3. What is the advantage of moving from a measured to a latent variable approach? What might happen to the estimates of effects with this transition?
4. Find an article in your area of interest that uses latent variable structural equation modeling (it may be referred to as structural equation modeling or covariance structures analysis). Read the article. Do the authors discuss reasons for using latent over measured variables? Do they link latent variables with reliability and validity? How do they label the disturbances? The error and unique variances of the measured variables?

Notes

1 Some of these examples are actually systematic errors rather than random errors and are thus not considered unreliability. I include them because I want you to consider the errors that can be included in even such a straightforward item.

2 The results are equivalent only because the example is so simple. With more items and multiple factors, the results of a confirmatory analysis will be different from those of an "exploratory" factor analysis (from SPSS), and even the results of an exploratory analysis will differ depending on the method used and the assumptions made. The example is useful for heuristic purposes, however, as a conceptual illustration of what factor analysis is.