Best Practices in Calculating Severe Discrepancies Between Expected and Actual Academic Achievement Scores: A Step-by-Step Tutorial

Jim Wright, Syracuse City Schools

Introduction. When diagnosing learning disabilities in school-age children, school psychologists typically look for a significant gap between the student's score on an aptitude, or cognitive, measure and (lower) performance on academic achievement testing. Indeed, the New York State Education Department states in its current Part 200 regulations governing special education services that "a student who exhibits a discrepancy of 50 percent or more between expected achievement and actual achievement determined on an individual basis shall be deemed to have a learning disability."

At present, schools use a variety of statistical and other formulas to determine whether a student has a severe discrepancy between expected and actual school achievement. This diversity of methods for identifying severe discrepancies makes it likely that evaluators in different school districts apply differing criteria to diagnose learning disabilities (Ross, 1992). A consensus has emerged in the research literature, though, about what methods comprise 'best practices' in calculating significant discrepancies between IQ and achievement test scores (Bennett & Clarizio, 1988; Reynolds, 1985): (1) test comparisons should be made using standardized scores (based on student age) rather than age- or grade equivalents or percentile rankings; (2) regression procedures should be used to take into account the partial correlation of IQ and achievement measures, and (3) score analyses should incorporate test-reliability data for each of the measures being compared (to control for score differences that can be traced to the tests' measurement characteristics rather than to the ability or skills of the person taking them).

Until recently, the complexity of the statistical calculations involved prevented many school psychologists from using those procedures most widely supported by researchers to compute severe discrepancies. Now, though, clinicians can use an Internet application, the Test Score Discrepancy Analyzer 2.0 (TSA2), to compute IQ-Achievement discrepancies (available at http://www.interventioncentral.org/tools.shtml). Originally developed as a tool for psychologists from one urban school district (Syracuse, NY), the program is being used increasingly by visitors from other school districts in New York and other states as well.

The TSA2 incorporates 'best practice' guidelines for statistical comparison of score discrepancies first recommended by the Special Education Programs Work Group on Measurement Issues in the Assessment of Learning Disabilities (Reynolds, 1985). Presented here is a tutorial that provides a detailed analysis of the statistical procedures clinicians can use to compute a discrepancy analysis of student IQ and achievement scores. Each step in this explanation provides the reader with a rationale for what must be accomplished and the computational formulas to be used. The tutorial also uses sample test data from a hypothetical student to illustrate the statistical operations required. The tutorial is based largely upon the work of Reynolds (1985). Most of the statistical formulas and notation appearing in this discussion are taken directly from Bennett & Clarizio (1988), whose article compares several score discrepancy formulas. Those wanting more information about test discrepancy issues are strongly encouraged to read Evans' (1990) article. Designed for the general reader, it presents an excellent and very accessible overview of the purpose and major stages of test discrepancy analysis. I also recommend Dumont and Willis' (1999) succinct and helpful web-based tutorial on calculating severe discrepancies.

Step 1: Assemble the Necessary Test Statistics

What We Need to Accomplish

To compute the size of discrepancy between an intelligence and academic achievement test, we will first need to collect basic statistical information about each test. For both the IQ and achievement measures, we will need to know the:

- test mean
- test standard deviation
- internal consistency reliability coefficient for the student's age
- student's actual test score

We also need to know the correlation between the IQ and achievement tests. (If there is no information available about the shared correlation between these tests, this value is estimated (Reynolds, 1985).

**Discrepancy Example**

In our example, we will compute a discrepancy using the following statistics from IQ and achievement tests:

<table>
<thead>
<tr>
<th>IQ Test Data:</th>
<th>Achievement Test Data:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Mean: 100</td>
<td>Test Mean: 100</td>
</tr>
<tr>
<td>Test SD: 15</td>
<td>Test SD: 15</td>
</tr>
<tr>
<td>Internal Consistency</td>
<td>Internal Consistency</td>
</tr>
<tr>
<td>Reliability Coefficient:</td>
<td>.96</td>
</tr>
<tr>
<td>Student's Test Score:</td>
<td>Student's Test Score:</td>
</tr>
<tr>
<td>98</td>
<td>82</td>
</tr>
</tbody>
</table>

Shared correlation between IQ and achievement tests = .715

---

**Step 2: Convert IQ & Achievement Scores to Z-Scores**

**What We Need to Accomplish**

Often we may wish to analyze the discrepancy between two tests that have different means and standard deviations. Our first step in the analysis, then, is to standardize test scores by converting them to z-scores. A z-score expresses a test score in *standard deviation units*. If a child attained a score of 115 on a test with a mean of 100 and a standard deviation of 15, for example, we can think of her as having performed one standard deviation above the mean. The z-score equivalent of 115 would be 1.0 (1 SD above the mean). When two tests with different mean and standard deviations have been converted to z-scores, we can compare them directly.

**Statistical Notation & Computational Formulas**

To transform a test score to a z-score, use the following formula:

\[ Z_X = \frac{X - \bar{X}}{\sigma_X} \]

In this formula:

- \(X\) = the student's score on the test
- \(\bar{X}\) = the test mean
- \(\sigma_X\) = the test standard deviation

**Discrepancy Example**

When we convert the IQ and achievement measures in our example, we get the following z-score equivalents:

\[ Z_{IQ} = \frac{(98-100)}{15} = -0.133 \]
\[ Z_{ACH} = \frac{(82-100)}{15} = -1.2 \]

Note: We can always convert z-score values back to standard test scores by using this formula:

\[ \text{Test Score} = (z\text{-score} \times \text{test SD}) + \text{test mean} \]

Here is an example of how we would convert our IQ test z-score back to a standard test score:
Step 3: Conduct a Significance Test of the IQ/Achievement Score Gap

What We Need to Accomplish

Before we go any further, we want to conduct a simple significance test of the gap between the student scores on the IQ and achievement measures (Figure 1). After all, it makes little sense (and can actually be misleading) to run discrepancy analyses on sets of scores whose difference may simply be a fluke.

For this test, we convert both IQ and achievement scores to z-scores so that we can compare them directly. We then complete a statistical significance test to answer the question: Is the score gap between the IQ and achievement measures greater than chance alone can reasonably account for? If the score gap is greater than chance alone can explain (i.e., is found to be significant), we go on to complete the remainder of the statistical analysis outlined below. If the score gap does not reach the threshold of significance, we classify the score gap as "Non-significant" and stop the analysis here.

Statistical Notation & Computational Formulas

We use the following formula to compute a significance value in z-score units for the IQ / Achievement discrepancy:

$$ Z = \frac{Z_x - Z_y}{\sqrt{2(1-\gamma_{xx} - \gamma_{yy})}} $$

In this formula:

- $Z$ = the magnitude of difference between IQ & ACH tests (expressed in standard-deviation units)
- $Z_x$ = the student's IQ test score in z-score units
- $Z_y$ = the student's achievement test score in z-score units
- $\gamma_{xx}$ = the internal consistency reliability coefficient for the IQ test
- $\gamma_{yy}$ = the internal consistency reliability coefficient for the achievement test

The computational formula used here is taken from Reynolds, 1985 (p.459). To be conservative, we are running the significance test as a two-tailed test. We set a confidence level of .95, which in a one-tailed test corresponds to a cut-off value (in z-score units) of 1.65 (Reynolds, 1985, p.459). If the value that we get from the IQ/Achievement significance formula exceeds this critical cut-off, we continue with the discrepancy analysis. If it does not, we stop our analysis here.

It is worth pointing out that this formula is set up so that, as test reliabilities decrease, there is a reduced likelihood that a gap will be found to be significant.

Discrepancy Example

When we run a significance test using our own values, adopting 1.65 as our cut-off, we get the following results:

$$ (-.133 - (-1.2)) / (2(1-.96)-.95))^{1/2} = 3.556 $$
Because our calculations yield a value above the 1.65 cut-off, our IQ / achievement score gap is considered "significant." The simple fact that significance was found, however, indicates simply that the gap between these scores is "real" and not due simply to chance. We must do further analysis to determine whether this score gap can be considered severe.

---

**Step 4: Compute an Estimated Student Achievement Score**

**What We Need to Accomplish**

In this step, we compute an expected achievement score for the student that takes into account the statistical concept of "regression to the mean". As Evans (1990) points out, when working with test statistics, we can visualize the concept of 'regression to the mean' by thinking of the achievement test score as being tugged at by two opposite but powerful attractors As the correlation between the IQ and achievement tests becomes higher, the estimated achievement score is 'pulled' from its own mean toward the IQ value. That is, as the correlation between two tests increases, we can use that shared correlation to predict with increasing confidence the estimated achievement score simply by knowing the IQ score. On the other hand, as the correlation between the IQ and achievement tests becomes lower, the estimated achievement score is 'pulled' back toward the mean value of the achievement test. The achievement test mean, rather than the IQ score, becomes the more powerful predictor of how the student will score on the achievement test. In fact, IQ and achievement tests are imperfectly correlated. When we compute an estimated achievement score, this score takes into account the twin influences of the IQ test score and the achievement test mean. The degree of correlation between IQ and achievement tests determines how much each source will shape the final estimated achievement test value.

**Statistical Notation & Computational Formulas**

To compute the estimated achievement score, you multiply the student's IQ z-score by the shared correlation between the IQ and achievement tests.

\[
Z'\hat{y} = \gamma_{xy}Z_x
\]

In this formula:

- \(Z'\hat{y}\) = the z-score value of the student's estimated achievement score
- \(\gamma_{xy}\) = the correlation between the IQ and achievement tests
- \(Z_x\) = the z-score value of the IQ test

**Discrepancy Example**

To compute an estimated achievement test score, we first plug our test values into the formula:

\[
Z'\hat{y} = (.715)*(-0.133) = -0.0944
\]

Then we convert this z-score to a standard achievement test score:

**Estimated achievement score** = (-0.0944*15)+100 = 98.58

---

**Step 5: Calculate the Difference Between Expected and Actual Achievement Scores**

**What We Need to Accomplish**

In this step, we will calculate the size of the gap between the expected and actual student achievement scores.

**Statistical Notation & Computational Formulas**

To compute the difference between expected and actual student achievement scores, we use this formula:
In this formula:

\[ D_{\hat{y}} - z_y = Z_{\hat{y}} - Z_y \]

- \[ D_{\hat{y}} - z_y \] = difference between expected and actual achievement scores (expressed in z-score units)
- \( Z_{\hat{y}} \) = the z-score value of the estimated achievement score
- \( Z_y \) = the z-score value of the actual achievement score

**Discrepancy Example**

Using our sample test values, we find that:

\( Z_{\hat{y}} = -0.094 \)

\( Z_y = -1.2 \)

\[ D_{\hat{y}} - z_y = (-0.094) - (-1.2) = 1.106 \]

---

**Step 6: Calculate the Magnitude of the Gap Between Expected and Actual Achievement Scores**

**What We Need to Accomplish**

We must now evaluate the magnitude of the gap between the student's expected and actual achievement scores to determine if the size of this gap is larger than chance alone can reasonably explain. The range of possible expected/actual achievement gaps is assumed to be normally distributed a mean of 0 (meaning that the student's estimated and actual achievement scores are identical). We divide the discrepancy value (calculated in Step 5) by the standard deviation of expected/actual score differences. The resulting value will tell us how far this particular expected/actual achievement score gap lies from the mean for such a gap.

**Statistical Notation & Computational Formulas**

Our computation goal is to convert our simple difference between estimated and actual achievement scores into a z-score. The \( Z_{\text{diff}} \) score will tell us how many standard deviations our difference value falls from the mean for such values (which is the student's predicted achievement score).

\[ Z_{\text{diff}} = \frac{D_{\hat{y}} - z_y}{\sqrt{1 - \gamma_{xy}^2}} \]

In this formula:

- \( Z_{\text{diff}} \) = the number of standard deviations (in z-score units) that the actual expected/actual score gap lies from the mean for such score gaps
- \( D_{\hat{y}} - z_y \) = difference between expected and actual achievement scores (expressed in z-score units)
- \( 1 - \gamma_{xy}^2 \) = the standard deviation of differences between expected and actual achievement scores

**Discrepancy Example**
When we calculate the magnitude of gap between the student's predicted and actual achievement scores, we find that:

\[
D = z_y - \bar{y} \\
(1 - \rho_{xy}^2)^{1/2} = (1-(.715)^2)^{1/2} = .699 \\
Z_{diff} = 1.106/.699 = 1.58
\]

Essentially we have created a new distribution in this step. The mean for the distribution is the student's predicted achievement score of 98.58. The standard deviation of this new distribution of discrepancies between estimated and actual achievement scores in test-score units is about 10.5 (z-score SD of .699 * test SD of 15 = 10.48). Figure 2 shows the new distribution. (Note that the student's actual achievement score of 82 falls well short of 2 SDs from the mean.)

**Figure 2: Comparison of Estimated & Actual Achievement Scores**

<table>
<thead>
<tr>
<th>Actual ACH</th>
<th>Estimated ACH</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>98.58</td>
</tr>
</tbody>
</table>

**Step 7: Adjust the Severe Discrepancy Cut-Off to Account for Test Unreliability**

*What We Need to Accomplish*

Had we stopped at Step 6, we would probably find that the student in our example does not have a severe discrepancy between expected and actual achievement scores. If we follow the advice of Reynolds (1985) and set a reasonable discrepancy cut-off score of 1.96 (two-tailed test; \( p = .05 \)), the student's \( Z_{diff} \) score of 1.58 would not meet this cut-off.

We have, however, one final important calculation to make before we can definitively decide whether a student's actual achievement score is severely discrepant. Some of the variation of test scores is due to unreliability (measurement error) within the test itself. It is important to adjust our discrepancy cut-off score upward to take into account measurement error. If we fail to do so, some students whose hypothetical "true score" on an achievement test falls within the severely discrepant range will attain actual scores that, because of measurement error alone, do not quite reach the severe discrepancy cut-off. When the cut-off is adjusted to account for test unreliability, we increase our confidence that our cut-off score does not unfairly screen out students because of the imperfect measurement characteristics of the tests used (Reynolds, 1985).

*Statistical Notation & Computational Formulas*

We set an initial cut-off score of 1.96 (\( p = .05 \) for a two-tailed test). Then we complete the dizzyingly complex series of calculations below to adjust the cut-off score upward as needed to take into account test unreliabilities:

\[
Z_{mod} = z_a \cdot 1.65 (1 - \rho_{\hat{y} - y_t})^{1/2} \\
where \quad z_a = z_{.05} = 1.96
\]
**Discrepancy Example**

Here are the components that we will need from our sample test data to compute the revised \(Z_{\text{mod}}\) cut-off:

\[
\begin{align*}
Z_a &= 1.96 \\
\gamma_{xx} &= .96 \\
\gamma_{yy} &= .95 \\
\gamma_{xy} &= .715
\end{align*}
\]

When we put these values into the computational formula, we get:

\[
\hat{y} - y_i = .95 + (.96 \cdot .715^2) - 2(.715^2) / 1 - (.715^2) = .428 / .4959 = .856
\]

\[
Z_{\text{mod}} = 1.96 - (1.65 (1 - .856)^{1/2}) = (1.96 - .6208) = 1.334
\]

So in our example, we find that the student's \(z_{\text{diff}}\) score of 1.57 exceeds our adjusted \(Z_{\text{mod}}\) cut-off score of about 1.33. We should therefore regard the discrepancy found between the student's expected and actual achievement scores as both **significant** (Step 3) and **severe** (Step 6).

---

**Step 8: Translate the Adjusted Cut-Off Score into a 'Critical Score' in Achievement Test Units**

**What We Need to Accomplish**

We are just about done! Now all we need to do is to use the adjusted cut-off score that we came up with in Step 7 \((Z_{\text{mod}})\) to calculate the threshold critical achievement test score that signifies a severe discrepancy. (The formula below is taken from Reynolds (1990) p.20).

**Statistical Notation & Computational Formulas**

\[
\text{Critical ACH Score} = \text{Est. ACH Score} - (Z_{\text{mod}} \cdot \sigma_y \cdot (1 - \gamma_{xy}^2)^{1/2})
\]

In this formula:

\[
Z_{\text{mod}} = \text{the modified critical cut-off score calculated in Step 7}
\]
Discrepancy Example

Here are the components that we will need from our sample test data to compute the critical achievement score that signifies a severe discrepancy:

<table>
<thead>
<tr>
<th>Estimated Achievement Score</th>
<th>98.58 (Computed in Step 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z_{mod}$</td>
<td>1.334 (Computed in Step 7)</td>
</tr>
<tr>
<td>$\sigma_y$</td>
<td>15</td>
</tr>
<tr>
<td>$(1 - r_{xy}^2)^{1/2}$</td>
<td>.699 (Computed in Step 6)</td>
</tr>
</tbody>
</table>

When we put these values into the computational formula, we get:

$$\text{Critical Achievement Test Score} = 98.58 - (1.334 \times 15 \times .699) = 84.59$$

Because our student's actual achievement score of 82 falls below the critical threshold test-score value of 84.59, the student's score is considered to be severely discrepant.

References


