Journal of Educational and Behavioral Statistics
June 2008, Vol. 33, No. 2, pp. 204–229
DOI: 10.3102/1076998607302636

When Can Subscores Have Value?

Shelby J. Haberman
Educational Testing Service

In educational tests, subscores are often generated from a portion of the items in a larger test. Guidelines based on mean squared error are proposed to indicate whether subscores are worth reporting. Alternatives considered are direct reports of subscores, estimates of subscores based on total score, combined estimates based on subscores and total scores, and residual analysis of subscores. Applications are made to data from two testing programs.

Keywords: *reliability; Kelley's formula; mean squared error; regression; true score*

Testing programs face increasing customer demand for subscores to provide more detailed information about an examinee than is provided by a total score. This demand may reflect a desire for diagnostic information or a desire for more data to facilitate placement or admission decisions. The intuitive approach to subscores is to divide the test items into separate sections that reflect different portions of the domain under study and to report scores for each separate section. For example, a writing test might include an essay and a multiple-choice section. The total score would be a sum of an essay subscore based entirely on the essay and a multiple-choice subscore based entirely on the multiple-choice section. The intuitive approach would involve a report that included the total score, the essay subscore, and the multiple-choice subscore. The intention would be that the essay portion provides a direct measure of writing proficiency, whereas the multiple-choice portion might be used to measure mastery of mechanics. This intuitive approach may not necessarily be the most informative one. For example, the essay subscore, which may be based on only one essay, may have only limited reliability. To overcome limited reliability of subscores, one may use regression to produce a linear combination of subscores that provides a more accurate estimate of the proficiency measured by the subscore than is provided by the intuitive subscore measure (Wainer et al., 2001; Yen, 1987).

However, a more important problem with the derivation of subscores from many existing educational tests is that there may be, as a matter of fact, relatively

The author thanks Neil Dorans, John Mazzeo, Sandip Sinharay, Alina von Davier, Matthias von Davier, Wendy Yen, Jinming Zhang, and anonymous referees for helpful discussions and comments. Results concerning the SAT I examination are derived from data provided by the College Board and are used with permission. Any opinions expressed in this article are those of the author and not necessarily of the Educational Testing Service.

little distinct information in the smaller subsets of items used for subscores that is not reflected in the total test score. When regressed subscores are used in such contexts, they are essentially replications of the total score and are uninformative. This problem reflects the nature of the test rather than the methodology used to develop a subscore. Tests not built to yield diagnostic information tend not to do so.

In this article, guidelines based on mean squared error (MSE) are proposed to indicate whether subscores are worth reporting for a particular test. Examples using data from two large-scale testing programs illustrate the results.

The proposed criterion for reporting of a subscore based on a portion of the items in a larger test is whether the subscore provides a more accurate measure of the construct it measures than is provided by the total score from the larger test. This standard for subscore reporting is readily handled by using the concept of true scores associated with classical test theory (Lord & Novick, 1968). Arguments are based on least squares and MSE. Section 1 provides the basic theory required for the analysis, whereas Section 2 considers estimation of required parameters. Computations are elementary and are based on reliability computations and elementary calculations of summary statistics. Section 3 considers some examples from testing programs at the Educational Testing Service. Section 4 provides some conclusions from the results of the analysis.

1. Mean Squared Errors for True Subscores

The basic criterion used in this article is the MSE associated with an approximation for the unobserved true subscore τ_X of an examinee associated with an observed subscore S_X based on a portion of the items from a larger assessment. Based on all items from the assessment, the examinee has an observed total score S_Z. MSE is employed to assess three basic strategies for approximation of the true subscore τ_X. In the first strategy, the observed subscore S_X is used to approximate τ_X. In the second strategy, the observed total score S_Z provides the approximation to the true subscore τ_X. In the third strategy, both the observed subscore S_X and the observed total score S_Z are employed to approximate τ_X. In all cases, linear approximations are employed.

To properly define the criterion of MSE used in this article, it is necessary to consider the notion of true score found in classical test theory and to provide basic conditions under which MSEs will be properly defined. As in classical test theory (Holland & Hoskens, 2003), the observed subscore S_X for an examinee randomly selected from a population of examinees and for a test form randomly selected from a collection of parallel test forms is a random variable with finite mean and variance that is decomposed into a random variable τ_X, termed the true subscore, that depends only on the examinee and a random variable $e_X = S_X - \tau_X$, termed the error of measurement of the subscore, that depends both on the examinee and on the test form. Thus, $S_X = \tau_X + e_X$. The true score τ_X is the expected value of the observed score S_X given the examinee, so that the observed subscore

S_X and the true subscore τ_X are random variables with the same expectation $E(S_X)$, the error of measurement e_X of the subscore has expectation $E(e_X) = 0$, the variance $\sigma^2(\tau_X)$ is finite, and the error variance $\sigma^2(e_X)$ of the subscore is finite. In addition, the true subscore τ_X and the error of measurement e_X of the subscore are uncorrelated (Holland & Hoskens, 2003). To avoid trivial cases, it is assumed that the variance $\sigma^2(\tau_X)$ of the true score τ_X and the error variance $\sigma^2(e_X)$ of the error of measurement e_X are both positive. The variance $\sigma^2(S_X) = \sigma^2(\tau_X) + \sigma^2(e_X)$ (Lord & Novick, 1968) of the observed subscore S_X is then positive, so that the observed subscore S_X has a reliability coefficient $\rho^2(S_X, \tau_X) = \sigma^2(\tau_X)/\sigma^2(S_X)$ equal to the square of the correlation $\rho(S_X, \tau_X)$ of S_X and τ_X. The reliability coefficient is positive and less than 1 (Lord & Novick, 1968).

The observed total score S_Z considered is also a random variable with finite mean and variance with a similar decomposition into a true total score τ_Z and an error of measurement e_Z. Again, S_Z is the total score for a randomly selected examinee from the same population of examinees considered for the observed subscore S_Z and for a randomly selected test form from the same population of parallel test forms used to define the observed subscore S_X. The true total score τ_Z is a random variable with finite mean $E(\tau_Z) = E(S_Z)$ and finite variance $\sigma^2(\tau_Z)$, and τ_Z is the conditional expected value of the observed subscore S_Z given the examinee. The error of measurement $e_Z = S_Z - \tau_Z$ of the total score is a random variable with finite mean 0 and with error variance $\sigma^2(e_Z)$. The true total score τ_Z and the error of measurement e_Z of the total score are uncorrelated.

To avoid trivial cases and to exploit classical test theory for composite tests (Lord & Novick, 1968), it is helpful to consider the observed remainder score $S_Y = S_Z - S_X$ associated with the items from the complete test that are not considered in the subscore. Thus, the total score S_Z is a composite of the subscore S_X and the remainder score S_Y. Under classical test theory, $S_Y = \tau_Y + e_Y$, where $\tau_Y = \tau_Z - \tau_X$ is the true remainder score and $e_Y = e_Z - e_X$ is the true error of measurement of the remainder score. The random variables S_Y, τ_Y, and e_Y all have finite means and variances, $E(S_Y) = E(\tau_Y) = E(S_Z) - E(S_Z)$, and $E(e_Y) = 0$. In addition, the true remainder score τ_Y and the error of measurement e_Y are uncorrelated, and the error of measurement e_X of the subscore and the error of measurement e_Y of the remainder score are also uncorrelated. As a consequence, the error variance $\sigma^2(e_Z)$ of the total score is the sum of the error variance $\sigma^2(e_X)$ of the subscore and the error variance $\sigma^2(e_Y)$ of the remainder score. In addition, the error e_Z of measurement of the total score and the true subscore τ_X are uncorrelated, and the true total score τ_Z and the error of measurement e_X of the subscore are also uncorrelated. To avoid trivial cases, it is assumed that the error variance $\sigma^2(e_Y)$ of the remainder score and the variance $\sigma^2(\tau_Y)$ of the true remainder score are both positive. Given the previous assumption that the error variance $\sigma^2(e_X)$ of the subscore is positive, then it follows that the error variance $\sigma^2(e_Z)$ of the total score is positive and the variance $\sigma^2(S_Z) = \sigma^2(\tau_Z) + \sigma^2(e_Z)$

of the total score is also positive. In addition, the true subscore τ_X and the remainder score τ_Y have a properly defined correlation $\rho(\tau_X, \tau_Y)$. It is also assumed that the variance $\sigma^2(\tau_Z)$ of the true total score τ_Z is positive, so that the true subscore τ_X and the true total score τ_Z have a defined correlation $\rho(\tau_X, \tau_Z)$ and the reliability coefficient $\rho^2(S_Z, \tau_Z) = \sigma^2(\tau_Z)/\sigma^2(S_Z)$ of the observed total score is greater than 0 and less than 1.

Given these basic results from classical test theory, MSE can be examined (Lord & Novick, 1968). If A is a random variable with finite variance used to approximate the true subscore τ_X, then the approximation error is $\tau_X - A$, and the MSE is the expectation

$$E([\tau_X - A]^2) = \sigma^2(\tau_X - A) + [E(\tau_X - A)]^2.$$

Because the MSE is in squared units, it is common to consider the root mean squared error (RMSE) $[E([\tau_X - A]^2)]^{1/2}$, for this measure is in the same units as the true subscore τ_X. For example, the intuitive approximation of the true subscore τ_X by the observed subscore S_X has approximation error $\tau_X - S_X = -e_X$, MSE equal to the error variance $\sigma^2(e_X)$ of the subscore S_X, and RMSE equal to the standard error of measurement $\sigma(e_X)$ of the subscore. The trivial approximation of the true subscore τ_X by the constant predictor $E(S_X)$ has approximation error $\tau_X - E(S_X)$, MSE equal to the variance $\sigma^2(\tau_X)$ of the true score τ_X, and RMSE equal to the standard deviation $\sigma(\tau_X)$ of the true score τ_X. Thus, the trivial approximation $E(S_X)$ to the true subscore τ_X is superior in terms of MSE to the intuitive approximation S_X if, and only if, the reliability coefficient $\rho^2(S_X, \tau_X)$ is less than one half.

In this article, the trivial predictor $E(S_X)$ will be used as a basic standard to judge other predictors. For a predictor A of τ_X with finite mean and variance, the proportional reduction $\Psi(\tau_X|A)$ of MSE of the approximation of τ_X by A relative to the approximation by $E(S_X)$ is then the relative decrease in MSE from use of A rather than $E(S_X)$. Thus,

$$\Psi(\tau_X|A) = 1 - E([\tau_X - A]^2)/\sigma^2(\tau_X),$$

so that $\Psi(\tau_X|E(S_X)) = 0$ and

$$\Psi(\tau_X|S_X) = 2 - 1/\rho^2(S_X, \tau_X).$$

In general, the proportional reduction in MSE $\Psi(\tau_X|A)$ is never greater than 1, and $\Psi(\tau_X|A)$ can be negative, as is the case for $\Psi(\tau_X|S_X)$ if the reliability $\rho^2(S_X, \tau_X)$ is less than one half.

If the approximation A has a positive standard deviation $\sigma(A)$, then it is a straightforward matter to employ linear regression to obtain an attractive alternative approximation. Let $\beta(\tau_X|A) = \rho(A, \tau_X)\sigma(\tau_X)/\sigma(A)$ denote the regression coefficient of the true subscore τ_X on A (Lord & Novick, 1968). Then, the linear regression

$$L(\tau_X|A) = E(S_X) + \beta(\tau_X|A)[A - E(A)]$$

of τ_X on A has an error equal to the residual

$$R(\tau_X|A) = \tau_X - L(\tau_X|A)$$

of the true subscore τ_X with respect to A (Lord & Novick, 1968), the expectation $E(R(\tau_X|A))$ is 0, the MSE from the approximation of the true subscore τ_X by $L(\tau_X|A)$ is

$$E([R(\tau_X|A)]^2) = \sigma^2(R(\tau_X|A)) = \sigma^2(\tau_X)[1 - \rho^2(A, \tau_X)],$$

the RMSE $\sigma(R(\tau_X|A))$ is $\sigma(\tau_X|A)) = \sigma(\tau_X)[1 - \rho^2(A, \tau_X)]^{1/2}$, and the proportional reduction in MSE is

$$\Psi(\tau_X|L(\tau_X|A)) = 1 - \rho^2(A, \tau_X) \geq 0.$$

For any real constants a and b, $\sigma^2(R(\tau_X|A)) \leq E([\tau_X - (a + bA)]^2)$, with equality if, and only if, $a = E(\tau_X) - \beta(\tau_X|A)$ and $b = \beta(\tau_X|A)$, so that $a + bA = L(\tau_X|A)$ (Lord & Novick, 1968). In particular, $\Psi(\tau_X|L(\tau_X|A))$ is less than 1 if τ_X and A are correlated, and $\Psi(\tau_X|L(\tau_X|A))$ exceeds $\Psi(\tau_X|A)$ unless the linear regression $L(\tau_X|A)$ of τ_X on A is equal to A.

In the following subsections, a series of applications of linear regression is considered to compare use of the observed subscore S_X, the observed total score S_Z, and the combination of S_X and S_Z for approximation of the true subscore τ_X. The criterion of MSE, or equivalently, the criterion of proportional reduction in MSE, then provides a basis for decisions on use of the observed subscore S_X, the observed total score S_X, or a combination of the observed subscore S_X and observed total score S_X for approximation of the true subscore τ_X. The relative value of the observed subscore is found to decrease as the reliability of the observed subscore decreases, the reliability of the observed total score increases, and the correlation $\rho(\tau_X, \tau_Z)$ of the true subscore τ_X and the true total score τ_Z increases.

1.1. Kelley's Formula

To begin, consider use of the linear regression $L(\tau_X|S_X)$ for approximation of the true subscore τ_X. This approach, as is well known, leads to Kelley's formula (Kelley, 1947)

$$L(\tau_X|S_X) = E(S_X) + \rho^2(S_X, \tau_X)[S_X - E(S_X)]$$

for the linear regression of the true subscore τ_X on the observed subscore S_X. For the linear regression $L(\tau_X|S_X)$, the approximation error is the residual $R(\tau_X|S_X) = \tau_X - L(\tau_X|S_X)$ of τ_X with respect to the observed subscore S_X, and the MSE $E([R(\tau_X|S_X)]^2)$ is the residual variance

$$\sigma^2(R(\tau_X|S_X)) = \sigma^2(\tau_X)[1 - \rho^2(S_X, \tau_X)] = \sigma^2(e_X)\rho^2(S_X, \tau_X)$$

of the linear regression of the true subscore τ_X on the observed subscore S_X. The proportional reduction $\Psi(\tau_X|L(\tau_X|A))$ in MSE reduces to the reliability $\rho^2(S_X, \tau_X)$ of the observed subscore S_X. It is always the case that $\Psi(\tau_X|L(\tau_X|A))$ exceeds $\Psi(\tau_X|S_X)$ and $\Psi(\tau_X|E(S_X)) = 0$.

Practical use of Kelley's formula requires estimation of $\beta(\tau_X|S_X)$ and $E(S_X)$ from sample data. This issue is considered in Section 2.1. In the empirical cases examined in Section 3, sample sizes are sufficiently large that problems of estimation from samples have a negligible effect in the analysis. Sample sizes will also be large enough so that estimation will have negligible effects on other approximations discussed in this section.

1.2. Approximation by Observed Total Score

A simple alternative to the use of Kelley's formula for the subscore is to approximate the true subscore τ_X by the linear regression $L(\tau_X|S_Z)$ of τ_X on the observed total score S_Z (Holland & Hoskens, 2003). It is important to note that this linear approximation based on the observed total score can be better in some instances than the linear approximation based on the observed subscore, even though the observed subscore is a more direct measure of the true subscore than is the observed total score. As is evident from the argument in this section, this possible advantage of the observed total score can arise if the reliability of the observed total score is higher than is the reliability of the observed subscore and if there is a sufficiently high correlation $\rho(\tau_X, \tau_Z)$ of the true subscore τ_X and the true total score τ_Z.

The linear regression

$$L(\tau_X|S_Z) = E(\tau_X) + \beta(\tau_X|S_Z)[S_Z - E(S_Z)]$$

of the true subscore τ_X on the observed total score S_Z has regression coefficient

$$\beta(\tau_X|S_Z) = \rho(\tau_X, S_Z)\sigma(\tau_X)/\sigma(S_Z),$$

where the standard deviation of the true subscore τ_X is

$$\sigma(\tau_X) = \sigma(S_X)\rho(S_X, \tau_X)$$

and the correlation of the true subscore τ_X and the observed total score S_Z satisfies

$$\rho(\tau_X, S_Z) = \rho(\tau_X, \tau_Z)\rho(S_Z, \tau_Z)$$

(Lord & Novick, 1968). The MSE $\sigma^2(R(\tau_X|S_Z))$ for approximation of the true subscore τ_X by the observed total score S_Z is then $\sigma^2(\tau_X)[1-\rho^2(\tau_X, S_Z)]$, and the proportional reduction in MSE $\Psi(\tau_X|L(\tau_X|S_Z))$ is $\rho^2(\tau_X, S_Z) = \rho^2(\tau_X, \tau_Z)\rho^2(S_Z, \tau_Z)$. Note that the proportional reduction $\Psi(\tau_X|L(\tau_X|S_Z))$ does not exceed the reliability coefficient $\rho^2(S_Z, \tau_Z)$ of the observed total score.

If the MSE $\sigma^2(R(\tau_X|S_Z))$ from the linear regression of the true subscore τ_X on the observed total score S_Z is less than the MSE $\sigma^2(R(\tau_X|S_X))$ from the linear regression of the true subscore τ_X on the observed subscore S_X, then use of the observed subscore S_X by itself is very difficult to justify for estimation of the true score τ_X, for the true subscore τ_X in this instance is better approximated by use of the linear regression on the observed total score S_Z than by use of the linear regression on the observed subscore S_X. The formulas for $\sigma^2(R(\tau_X|S_X))$ and $\sigma^2(R(\tau_X|S_Z))$ imply that the MSE $\sigma^2(R(\tau_X|S_Z))$ from the linear regression on the observed total score S_Z is less than the MSE $\sigma^2(R(\tau_X|S_X))$ from the regression on the observed subscore S_X if, and only if, the product $\rho^2(S_Z, \tau_Z)$ $\rho^2(\tau_X, \tau_Z)$ of the reliability coefficient of S_Z and the squared correlation of the true subscore τ_X and the true total score τ_Z exceeds the reliability coefficient $\rho^2(S_X, \tau_X)$ of the observed subscore S_X. Thus, use of the observed total score S_Z rather than the observed subscore S_X is increasingly favored as the reliability $\rho^2(S_Z, \tau_Z)$ of the total score S_Z increases, the correlation $\rho(\tau_X, \tau_Z)$ of true subscore τ_X and true total score τ_Z increases, and the reliability $\rho^2(S_X, \tau_X)$ of the subscore S_X decreases. The observed total score S_Z can only be a better predictor of the true subscore τ_X if the observed total score S_Z has higher reliability than the observed subscore S_X. In this case, use of the observed total score is favored if the correlation $\rho(\tau_X, \tau_Z)$ of the true scores is sufficiently high.

In Section 2.2, estimation procedures are developed to permit practical use of the linear regression approximation $L(\tau_X|S_Z)$ to the true subscore τ_X.

1.3. Approximation by Both Observed Subscore and Observed Total Score

The linear regression $L(\tau_X|S_X, S_Z)$ of the true subscore τ_X on both the observed subscore S_X and the observed total score S_Z may be used to approximate the true subscore τ_X. As an approximation to τ_X, $L(\tau_X|S_X, S_Z)$ has a MSE $\sigma(R(\tau_X|S_X, S_Z))$ no greater than the minimum of the MSE $\sigma^2(R(\tau_X|S_X))$ from the linear regression on τ_X on the observed subscore S_X and the MSE based on the observed total score (Wainer et al., 2001). The practical question is whether the improvement from joint use of the observed subscore and observed total score is substantial or negligible.

Unique definition of the linear regression $L(\tau_X|S_X, S_Z)$ requires that the correlation $\rho(S_X, S_Z)$ of the observed subscore S_X and the observed total score S_Z must satisfy the constraint that $|\rho(S_X, S_Z)| < 1$. In the appendix, it is shown that this constraint on the correlation $\rho(S_X, S_Z)$ is always satisfied.

The linear regression

$$L(\tau_X|S_X, S_Z) = E(S_X) + \beta(\tau_X|S_X \cdot S_Z)[S_X - E(S_X)] + \beta(\tau_X|S_Z \cdot S_X)[S_Z - E(S_Z)]$$

of the true subscore τ_X on the observed subscore S_X and the observed total score S_Z depends on the partial regression coefficient

$$\beta(\tau_X|S_X \cdot S_Z) = \beta(R(\tau_X|S_Z)|R(S_X|S_Z)) = \frac{\sigma(\tau_X)[\rho(S_X, \tau_X) - \rho(\tau_X, S_Z)\rho(S_X, S_Z)]}{\sigma(S_X)[1 - \rho^2(S_X, S_Z)]}$$

of the true subscore τ_X on the observed subscore S_X given the observed total score S_Z and on the partial regression coefficient

$$\beta(\tau_X|S_Z \cdot S_X) = \beta(R(\tau_X|S_X)|R(S_Z|S_X)) = \frac{\sigma(\tau_X)[\rho(\tau_X, S_Z) - \rho(S_X, \tau_X)\rho(S_X, S_Z)]}{\sigma(S_Z)[1 - \rho^2(S_X, S_Z)]}$$

of the true subscore τ_X on the observed total score S_Z given the observed subscore S_X (Lord & Novick, 1968). In the application of the formulas, recall that $\rho(\tau_X, S_Z) = \rho(\tau_X, \tau_Z)\rho(S_Z, \tau_Z)$ and $\sigma(\tau_X) = \sigma(S_X)\rho(S_X, \tau_X)$. The approximation error is $R(\tau_X|S_X, S_Z) = \tau_X - L(\tau_X|S_X, S_Z)$, the expectation $E(R(\tau_X|S_X, S_Z))$ is 0. For any real constants a, b, and c, the MSE $\sigma^2(R(\tau_X|S_X, S_Z))$ for approximation of the true subscore τ_X by the linear regression $L(\tau_X|S_X, S_Z)$ satisfies $\sigma^2(R(\tau_X|S_X, S_Z)) \leq E([\tau_X - (a_b S_X + cS_Z)]^2)$, with equality if, and only if, $a + bS_X + cS_Z = L(\tau_X|S_X, S_Z)$, so that $b = \beta(\tau_X|S_X \cdot S_Z)$, $c = \beta(\tau_X|S_Z \cdot S_X)$, and $a = (1 - b)E(S_X) - cE(S_Z)$.

The linear regression $L(\tau_X|S_X, S_Z)$ of the true subscore τ_X on the observed subscore S_X and the observed total score S_Z reduces to the Kelley approximation $L(\tau_X|S_X)$ based on the observed subscore S_X if $\rho(S_Z, \tau_Z)\rho(\tau_X, \tau_Z) = \rho(S_X, \tau_X)\rho(S_X, S_Z)$. If $\rho(S_X, \tau_X) = \rho(\tau_X, \tau_Z)\rho(S_Z, \tau_Z)\rho(S_X, S_Z)$, then the linear regression $L(\tau_X|S_X, S_Z)$ reduces to the linear regression $L(\tau_X|S_Z)$ of the true subscore on the observed total score S_Z.

Formulas for MSE rely on the partial correlation coefficient

$$\rho(S_X, \tau_X \cdot S_Z) = \rho(R(\tau_X|S_Z), R(S_X|S_Z)) = \frac{\rho(S_X, \tau_X) - \rho(\tau_X, S_Z)\rho(S_X, S_Z)}{[1 - \rho^2(\tau_X, S_Z)]^{1/2}[1 - \rho^2(S_X, S_Z)]^{1/2}}$$

of the true subscore τ_X and the observed subscore S_X given the observed total score S_Z and the partial correlation coefficient

$$\rho(\tau_X, S_Z \cdot S_X) = \rho(R(\tau_X|S_X), R(S_Z|S_X)) = \frac{\rho(\tau_X, S_Z) - \rho(\tau_X, S_X)\rho(S_X, S_Z)}{[1 - \rho^2(S_X, \tau_X)]^{1/2}[1 - \rho^2(S_X, S_Z)]^{1/2}}$$

of the true subscore τ_X and the observed total score S_Z given the observed subscore S_X (Lord & Novick, 1968). Again, recall that $\rho(\tau_X, S_Y) = \rho(\tau_X, \tau_Z)$ $\rho(S_Z, \tau_Z)$. The MSE $\sigma^2(R(\tau_X|S_X, S_Z))$ can be expressed in terms of the previous formulas for the MSE $\sigma^2(R(\tau_X|S_X))$ from approximation of the true subscore τ_X by use of Kelley's formula and the MSE $\sigma^2(R(\tau_X|S_Z))$ from approximation of the true subscore by use of the linear regression $L(\tau_X|S_Z)$ on the observed total score. One has

$$\sigma^2(R(\tau_X|S_X, S_Z)) = \sigma^2(R(\tau_X|S_X))[1 - \rho^2(\tau_X, S_Z \cdot S_X)] = \sigma^2(R(\tau_X|S_Z))[1 - \rho^2(\tau_X, S_X \cdot S_Z)],$$

so that the MSE $\sigma^2(R(\tau_X|S_X, S_Z))$ from the linear regression of the true subscore τ_X on both the observed subscore S_X and the observed total score S_Z cannot exceed either the MSE $\sigma^2(R(\tau_X|S_X))$ from the linear regression of the true subscore τ_X on the observed subscore S_X or the MSE $\sigma^2(R(\tau_X|S_Z))$ from the linear regression of the true subscore τ_X on the observed subscore S_Z. As is evident from the possibility that $L(\tau_X|S_X, S_Z)$ is either $L(\tau_X|S_X)$ or $L(\tau_X|S_Z)$, it is not necessarily the case that the MSE $\sigma^2(R(\tau_X|S_X, S_Z))$ from the linear regression on both the observed subscore S_X and the observed total score S_Z is less than the minimum of the MSEs $\sigma^2(R(\tau_X|S_X))$ and $\sigma^2(R(\tau_X|S_Z))$. The condition for equality of the two MSEs $\sigma^2(R(\tau_X|S_X, S_Z))$ and $\sigma^2(R(\tau_X|S_X))$ is the same as the condition for equality of the two linear regressions $L(\tau_X|S_X, S_Z)$ and $L(\tau_X|S_X)$. Similarly, the condition for equality of the two MSEs $\sigma^2(R(\tau_X|S_X, S_Z))$ and $\sigma^2(R(\tau_X|S_Z))$ is the same as the condition for equality of the two linear regressions $L(\tau_X|S_X, S_Z)$ and $L(\tau_X|S_Z)$.

The proportional reduction in MSE relative to the constant predictor $E(S_X)$ is

$$\begin{aligned}\Psi(\tau_X|L(\tau_X|S_X, S_Z)) &= 1 - [1 - \rho^2(S_X, \tau_X)][1 - \rho^2(\tau_X, S_Z \cdot S_X)] \\ &= 1 - [1 - \rho^2(\tau_X, S_Z)][1 - \rho^2(S_X, \tau_X \cdot S_Z)].\end{aligned}$$

Thus, appreciable gain from prediction with both the observed subscore S_X and the total subscore S_Z depends on partial correlation coefficients $\rho(S_X, \tau_X|S_Z)$ and $\rho(\tau_X, S_Z|S_X)$ that differ substantially from 0.

In Section 2.3, estimation procedures are developed to permit practical application of $L(\tau_X|S_X, S_Z)$.

Table 1 provides a summary of formulas for RMSE and proportional reduction of MSE for approximation of the true subscore τ_X.

TABLE 1
Formulas for Root Mean Squared Error (RMSE) and Proportional Reduction in Mean Squared Error (MSE) for Approximation of the True Subscore

Predictor	RMSE	Proportional Reduction in MSE
Subscore	$\sigma(\tau_X)[1-\rho^2(S_X,\tau_X)]^{1/2}$	$\rho^2(S_X,\tau_X)$
Total score	$\sigma(\tau_X)[1-\rho^2(S_Z,\tau_X)]^{1/2}$	$\rho^2(S_Z,\tau_X)$
Subscore, total score	$\sigma(\tau_X)[1-\rho^2(S_X,\tau_X)]^{1/2}$ $[1-\rho^2(S_Z,\tau_X\cdot S_X)]^{1/2}$	$1-[1-\rho^2(S_X,\tau_X)]$ $[1-\rho^2(S_Z,\tau_X\cdot S_X)]$

1.4. Approximation of the True Residual Subscore

An alternative approach to subscores considers the true residual subscore

$$R(\tau_X|\tau_Z)=\tau_X-E(S_X)-\beta(\tau_X|\tau_Z)[\tau_Z-E(S_Z)]$$

from the linear regression of the true subscore τ_X on the true total score τ_Z. Thus, $R(\tau_X|\tau_Z)$ may be regarded as the portion of the true subscore that is not predicted by the true total score. Indeed, $R(\tau_X|\tau_Z)$ and the true total score τ_Z are uncorrelated (Lord & Novick, 1968). A positive value of $R(\tau_X|\tau_Z)$ indicates that the true subscore τ_X is higher than expected from the total true score τ_Z, whereas a negative value of $R(\tau_X|\tau_Z)$ suggests a weaker true subscore τ_X than predicted by the true total score τ_Z.

The trivial approximation of $R(\tau_X|\tau_Y)$ is the constant predictor 0 that corresponds to a true subscore τ_X that is a linear function of the true total score τ_Z. Because $R(\tau_X|\tau_Z)$ has expected value 0, the MSE is then

$$\sigma^2(R(\tau_X|\tau_Z))=[1-\rho^2(\tau_X,\tau_Z)]\sigma^2(\tau_X), \tag{1}$$

so that the RMSE is

$$\sigma(R(\tau_X|\tau_Z))=[1-\rho^2(\tau_X,\tau_Z)]^{1/2}\sigma(\tau_X). \tag{2}$$

Note that the RMSE $\sigma(R(\tau_X|\tau_Z))$ is no greater than the RMSE $\sigma(\tau_X)$ from use of a constant predictor $E(S_X)$ for the true subscore τ_X. The RMSE is 0 if the true subscore τ_X is a linear function of the true total score τ_Z, so that $\rho^2(\tau_X,\tau_Z)=1$.

Alternatively, $R(\tau_X|\tau_Z)$ may be approximated by the linear regression

$$L(R(\tau_X|\tau_Z)|S_X,S_Z)=\beta(R(\tau_X|\tau_Z)|S_X\cdot S_Z)[S_X-E(S_X)]+\beta(R(\tau_X|\tau_Z)|S_Z\cdot S_X)[S_Z-E(S_Z)]$$

of the true residual $R(\tau_X|\tau_Z)$ on the observed subscore S_X and the observed total score S_Z. As shown in the appendix,

$$L(R(\tau_X|\tau_Z)|S_X, S_Z) = \frac{\sigma^2(R(\tau_X|\tau_Z))}{\sigma^2(R(S_X|S_Z))} R(S_X|S_Z),$$

where $R(S_X|S_Z)$ is the observed residual subscore. The multiplier $\sigma^2(R(\tau_X|\tau_Z)/\sigma^2(R(S_X|S_Z)$ of the observed residual subscore $R(S_X|S_Z)$ is nonnegative and less than 1. The residual of the residual subscore $R(\tau_X|\tau_Z)$ with respect to the observed subscore S_X and the observed total score S_Z is

$$R(R(\tau_X|\tau_Z)|S_X, S_Z) = R(\tau_X|\tau_Z) - L(R(\tau_X|\tau_Z)|S_X, S_Z),$$

and the MSE from approximation of the true residual subscore $R(\tau_X|\tau_Z)$ by the linear regression $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ of $R(\tau_X|\tau_Z)$ on the observed subscore S_X and the observed total score S_Z is the residual variance

$$\sigma^2(R(\tau_X|\tau_Z)|S_X, S_Z)) = \sigma^2(R(\tau_X|\tau_Z))[1 - \Psi(R(\tau_X|\tau_Z)|L(R(\tau_X|\tau_Z)|S_X, S_Z))],$$

where the proportional reduction in MSE

$$\Psi(R(\tau_X|\tau_Z)|L(R(\tau_X|\tau_Z)|S_X, S_Z)) = \frac{\sigma^2(R(\tau_X|\tau_Z))}{\sigma^2(R(S_X|S_Z))}$$

is less than 1 and at least 0.

Estimation procedures for this case are developed in Section 2.4.

2. Estimation of Parameters

In practice, the methods of Section 1 require parameter estimates derived by use of sampling. In this section, it is assumed that standard techniques have been employed to estimate the expectation $E(S_X)$ and standard deviation $\sigma(S_X)$ of the observed subscore S_X by respective estimates $\hat{E}(S_X)$ and $\hat{\sigma}(S_X)$ and to estimate the expectation $E(S_Z)$ and standard deviation $\sigma(S_Z)$ of the observed total score S_Z by respective estimates $\hat{E}(S_Z)$ and $\hat{\sigma}(S_Z)$. In addition, the correlation coefficient $\rho(S_X, S_Z)$ of the observed subscore S_X and the observed total score S_Z has been estimated by $\hat{\rho}(S_X, S_Z)$, the standard error of measurement $\sigma(e_X)$ of the observed subscore S_X has been estimated by $\hat{\sigma}(e_X)$, and the standard error of measurement $\sigma(e_Z)$ of the observed total score S_Z has been estimated by $\hat{\sigma}(e_Z)$ (Lord & Novick, 1968). In the analysis in this article, a test is divided into several subtests, and each subtest yields a subtest score. The total score is the sum of the subtest scores, and the subscore under study is one of these subtest scores. For each subtest, reliability coefficients, error variances, and standard errors of measurement are estimated by use of the KR-20 approach (Dressel, 1940; Kuder & Richardson, 1937). The error variance for the total score is then the sum of the error variances for the subtest scores. Given the error variance for the total test, the standard error of measurement for the total score is readily derived.

In Section 2.1, estimation for the Kelley formula $L(\tau_X|S_X)$ is developed. In Section 2.2, estimation for the linear regression $L(\tau_X|S_Z)$ of true subscore τ_X on observed total score S_Z is considered. Section 2.3 develops estimation for the linear regression $L(\tau_X|S_X, S_Z)$ of true subscore τ_X on both observed subscore S_X and observed total score S_Z. In Section 2.4, estimation procedures for analysis of the true residual subscore $R(\tau_X|\tau_Z)$ are provided.

2.1. Estimation for Approximation by Observed Subscore

In this case, only estimates for summary statistics related to the subscore S_X need be considered. The error variance $\sigma^2(e_X)$ is estimated by the square $\hat{\sigma}^2(e_X)$ of the estimated standard error of measurement $\hat{\sigma}(e_X)$ of the subscore S_X. The variance $\sigma^2(\tau_X)$ is then estimated by the difference $\hat{\sigma}^2(S_X) - \hat{\sigma}^2(e_X)$, and the reliability $\rho^2(S_X, \tau_X)$ is estimated by $\hat{\rho}^2(S_X, \tau_X) = \hat{\sigma}^2(\tau_X)/\hat{\sigma}^2(S_X)$. The constant approximation $E(S_X)$ is approximated by $\hat{E}(S_X)$. The RMSE $\sigma(\tau_X)$ associated with the constant $E(S_X)$ is then estimated by the square root $\hat{\sigma}(\tau_X)$ of the estimated variance $\hat{\sigma}^2(\tau_X)$ of the true subscore τ_X. The estimate of the linear regression $L(\tau_X|S_X)$ is then

$$\hat{L}(\tau_X|S_X) = \hat{E}(S_X) + \hat{\rho}^2(S_X, \tau_X)[S_X - \hat{E}(S_X)].$$

The MSE $\sigma^2(R(\tau_X|S_X))$ for the Kelley approximation is estimated by $\hat{\sigma}^2(R(\tau_X|S_X)) = \hat{\sigma}^2(\tau_X)[1 - \hat{\rho}^2(S_X, \tau_X)]$.

2.2. Estimation for Approximation by Observed Total Score

In addition to quantities already estimated in Section 2.1, several similar estimates are required for the total score. The error variance $\sigma^2(e_Z)$ of the total score is estimated by the square $\hat{\sigma}^2(e_Z)$ of the estimated standard deviation of measurement of the total score, the variance $\sigma^2(S_Z)$ of the observed total score S_Z is estimated by the square $\hat{\sigma}^2(S_Z)$ of the estimated standard deviation of S_Z, and the variance $\sigma(\tau_Z)$ of the total score is estimated by the difference $\hat{\sigma}^2(\tau_Z) = \hat{\sigma}^2(S_Z) - \hat{\sigma}^2(e_Z)$. The estimated reliability $\hat{\rho}^2(S_Z, \tau_Z)$ of the total score S_Z is then $\hat{\sigma}^2(\tau_Z)/\hat{\sigma}^2(S_Z)$. The estimated correlation $\hat{\rho}(S_X, \tau_X)$ of the observed subscore S_X and the true subscore τ_X is the square root of the estimated reliability $\hat{\rho}^2(S_X, \tau_X)$ of S_X, and the estimated correlation $\hat{\rho}(S_Z, \tau_Z)$ of the observed total score S_Z and the true total score τ_Z is the square root of the estimated reliability $\hat{\rho}^2(S_Z, \tau_Z)$ of S_Z.

Estimation of the correlation $\hat{\rho}(\tau_X, \tau_Z)$ of the true subscore τ_X and the true total score τ_Z requires a slightly more difficult argument because the true subscore τ_X is a component of the true total score τ_Z. Details are provided in the appendix. One finds that $\rho(\tau_X, \tau_Z)$ has estimate

$$\hat{\rho}(\tau_X, \tau_Z) = \frac{\hat{\rho}(S_X, S_Z)}{\hat{\rho}(S_X, \tau_X)\hat{\rho}(S_Z, \tau_Z)} - \frac{\hat{\sigma}^2(e_X)}{\hat{\sigma}(\tau_X)\hat{\sigma}(S_Z)}.$$

The estimated correlation $\hat{\rho}(\tau_X, S_Z)$ of the true subscore τ_X and the observed total score S_Z is then the product $\hat{\rho}(\tau_X, \tau_Z)\hat{\rho}(S_Z, \tau_Z)$. The linear regression

$$L(\tau_X|S_Z) = E(\tau_X) + \beta(\tau_X|S_Z)[S_Z - E(S_Z)]$$

of the true subscore τ_X on the observed total score S_Z is then estimated by

$$\hat{L}(\tau_X|S_Z) = \hat{E}(\tau_X) + \hat{\beta}(\tau_X|S_Z)[S_Z - \hat{E}(S_Z)],$$

where the estimated regression coefficient

$$\hat{\beta}(\tau_X|S_Z) = \hat{\rho}(\tau_X, S_Z)\hat{\sigma}(\tau_X)/\hat{\sigma}(S_Z).$$

The estimated MSE $\hat{\sigma}^2(R(\tau_X|S_Z))$ is then $\hat{\sigma}^2(\tau_X)[1 - \hat{\rho}^2(\tau_X, S_Z)]$, where $\hat{\rho}^2(\tau_X, S_Z)$ is the square of the estimated correlation $\hat{\rho}(\tau_X, S_Z)$ of τ_X and S_Z.

2.3. Estimation for Approximation by Both Observed Subscore and Observed Total Score

Given estimates derived in Sections 2.1 and 2.2, estimates required for application of Section 1.3 reduce to simple substitutions of previously derived estimates for the quantities they estimate. For example, the linear regression $L(\tau_X|S_X, S_Z)$ of the true subscore τ_X on the observed subscore S_X and the observed total score S_Z is estimated by

$$\hat{L}(\tau_X|S_X, S_Z) = \hat{E}(S_X) + \hat{\beta}(\tau_X|S_X \cdot S_Z)[S_X - \hat{E}(S_X)] + \hat{\beta}(\tau_X|S_Z \cdot S_X)[S_Z - \hat{E}(S_Z)],$$

where

$$\hat{\beta}(\tau_X|S_X \cdot S_Z) = \frac{\hat{\sigma}(\tau_X)[\hat{\rho}(S_X, \tau_X) - \hat{\rho}(\tau_X, S_Z)\hat{\rho}(S_X, S_Z)]}{\hat{\sigma}(S_X)[1 - \hat{\rho}^2(S_X, S_Z)]}$$

is the estimated partial regression coefficient of the true subscore τ_X on the observed subscore S_X given the observed total score S_Z and

$$\hat{\beta}(\tau_X|S_Z \cdot S_X) = \frac{\hat{\sigma}(\tau_X)[\hat{\rho}(\tau_X, S_Z) - \hat{\rho}(S_X, \tau_X)\hat{\rho}(S_X, S_Z)]}{\hat{\sigma}(S_Z)[1 - \hat{\rho}^2(S_X, S_Z)]}$$

is the estimated partial regression coefficient of the true subscore τ_X on the observed total score S_Z given the observed subscore S_X.

2.4. Estimation for Approximation of the True Residual Subscore

As in Section 2.3, substitution of estimates for quantities estimated is quite adequate for implementation of results of Section 1.4. The basic results to note are that the estimated variance $\hat{\sigma}^2(R(\tau_X|\tau_Z))$ for the residual true score $R(\tau_X|\tau_Z)$ is $[1-\hat{\rho}^2(\tau_X,\tau_Z)]\,\hat{\sigma}^2(\tau_X)$, and $\hat{\sigma}^2(R(S_X|S_Z)=\hat{\sigma}(S_X)[1-\hat{\rho}^2(S_X,S_Z)]$ is the estimated variance of the residual subscore $R(S_X|S_Z)$. Thus, the linear regression $L(R(\tau_X|\tau_Z)|S_X,S_Z)$ has estimate

$$\hat{L}(R(\tau_X|\tau_Z)|S_X,S_Z)=\frac{\hat{\sigma}^2(R(\tau_X|\tau_Z))}{\hat{\sigma}^2(R(S_X|S_Z))}R(S_X|S_Z).$$

The remaining estimation is straightforward.

3. Examples

To illustrate application of subscores, consider data from the October 2002 administration of the SAT I examination (Feigenbaum & Hammond, 2003). Results are summarized in Tables 2 through 7. Reference to Table 1 may be helpful in terms of formulas used. In these tables, Verbal I, Verbal II, and Verbal III refer to the three separate sections of the SAT verbal examination, which are interleaved with Math I, Math II, and Math III, the three separate sections of the SAT math examination. In Table 2, computations of standard errors of measurement for the complete verbal test, the complete math test, and for the sum of the verbal and math test are based on decomposition of the verbal test into subtests Verbal I to Verbal III and decomposition of the math test into subtests Math I to Math III. An alternative breakdown of the SAT verbal uses critical reading (CR), analogies (A), and sentence completion (SC). Similarly, an alternate decomposition of SAT math uses four-choice math multiple choice (Math 4c), five-choice multiple choice (Math 5c), and student-produced math responses (Math S). It should be noted that the SAT I examination has changed significantly since 2002. In addition, none of the section scores or subscores considered here were reported to institutions, although raw scores on scales CR, A, and SC were reported to examinees.

In these tables, for any given line, the observed subscore is S_X and the observed total score is S_Z. For example, let S_X be the subscore for critical reading, and let S_Z be the total score for the 78-item verbal test. Then, in Table 3, the standard error of measurement $\sigma(e_X)$ of the observed subscore S_X is estimated to be 3.4, whereas the RMSE $\sigma(R(\tau_X|S_X))$ from approximation of the true subscore τ_X by use of the Kelley approximation $L(\tau_X|S_X)$ is estimated to be 3.2. In this case, the observed subscore S_X from the first verbal section is clearly

TABLE 2
Summary Statistics for SAT Subscores

Subscore	Number of Items	Mean	Standard Deviation	Standard Error of Measurement	Total	Correlation With Total
Verbal I	36	14.9	7.3	2.9	Verbal	.95
Verbal II	30	12.9	6.3	2.8	Verbal	.94
Verbal III	12	5.2	3.4	1.8	Verbal	.85
CR	40	19.4	8.6	3.4	Verbal	.96
A	19	9.3	4.1	2.1	Verbal	.87
SC	19	10.6	4.4	2.1	Verbal	.90
Math I	25	13.8	6.3	2.3	Math	.95
Math II	25	13.2	5.5	2.3	Math	.95
Math III	10	5.5	2.5	1.5	Math	.82
Math 4c	15	6.9	4.2	0.9	Math	.88
Math 5c	35	15.8	8.4	2.1	Math	.98
Math S	10	4.4	2.4	0.7	Math	.85
Verbal	78	32.7	16.3	4.4	Total	.93
Math	60	26.9	14.1	3.6	Total	.91

Note: CR = critical reading; A = analogies; SC = sentence completion.

TABLE 3
Root Mean Squared Errors for Approximations for SAT True Subscores

Subscore	$\sigma(e_X)$	$\sigma(R(\tau_X\|S_X))$	$\sigma(R(\tau_X\|S_Z))$	$\sigma(R(\tau_X\|S_X, S_Z))$
Verbal I	2.9	2.7	2.1	2.0
Verbal II	2.8	2.5	1.6	1.6
Verbal III	1.8	1.5	1.1	1.0
CR	3.4	3.2	2.6	2.6
A	2.1	1.8	1.3	1.2
SC	2.1	1.8	1.3	1.3
Math I	2.3	2.1	1.7	1.6
Math II	2.3	2.1	1.5	1.4
Math III	1.5	1.2	0.7	0.7
Math 4c	1.9	1.6	1.0	1.0
Math 5c	2.7	2.6	2.1	2.1
Math S	1.2	1.1	0.7	0.7
Verbal	4.4	4.2	5.7	4.1
Math	3.6	3.4	5.4	3.3

Note: CR = critical reading; A = analogies; SC = sentence completion.

unsatisfactory relative to the linear regression $L(\tau_X|S_Z)$ of the true subscore τ_X from the first verbal section on the total verbal score S_Z, for the RMSE $\sigma(R(\tau_X|S_Z))$ is estimated to be 2.6, a somewhat smaller figure than is available

TABLE 4
Linear Regression Weights for Approximation for SAT True Subscores by Observed Subscores and Observed Total Scores

Subscore	$\beta(\tau_X \mid S_X \cdot S_Z)$	$\beta(\tau_X \mid S_Z \cdot S_X)$
Verbal I	0.15	0.34
Verbal II	−0.01	0.35
Verbal III	0.20	0.13
CR	0.17	0.38
A	0.18	0.17
SC	0.20	0.18
Math I	0.16	0.35
Math II	0.06	0.34
Math III	0.06	0.13
Math 4c	0.04	0.21
Math 5c	0.13	0.48
Math S	0.16	0.12
Verbal	0.71	0.13
Math	0.77	0.09

Note: CR = critical reading; A = analogies; SC = sentence completion.

TABLE 5
Proportional Reduction of Mean Squared Error Achieved by Approximations of SAT True Subscores

Subscore	Total score	$\Psi(\tau_X \mid L(\tau_X \mid S_X))$	$\Psi(\tau_X \mid L(\tau_X \mid S_Z))$	$\Psi(\tau_X \mid L(\tau_X \mid S_X, S_Z))$
Verbal I	Verbal	0.84	0.91	0.91
Verbal II	Verbal	0.80	0.92	0.92
Verbal III	Verbal	0.72	0.85	0.87
CR	Verbal	0.84	0.89	0.89
A	Verbal	0.74	0.87	0.88
SC	Verbal	0.78	0.88	0.89
Math I	Math	0.87	0.92	0.92
Math II	Math	0.83	0.92	0.92
Math III	Math	0.64	0.89	0.89
Math 4c	Math	0.72	0.90	0.90
Math 5c	Math	0.89	0.92	0.92
Math S	Math	0.73	0.89	0.90
Verbal	Total	0.91	0.85	0.92
Math	Total	0.92	0.82	0.93

Note: CR = critical reading; A = analogies; SC = sentence completion.

from the first verbal subscore S_X itself. Use of the linear regression $L(\tau_X \mid S_X, S_Z)$ of the true subscore τ_X on the observed subscore S_X and the observed total score S_Z yields only a slight reduction in RMSE, for $\sigma(R(\tau_X \mid S_X, S_Z))$ is also 2.6 to two

TABLE 6

Root Mean Squared Error for Approximation of SAT True Residual Subscores

Subscore	Total Score	$\sigma(R(\tau_X\|\tau_Z))$	$\sigma(R(R(\tau_X\|\tau_Z)\|S_X S_Z))$
Verbal I	Verbal	0.86	0.80
Verbal II	Verbal	0.40	0.39
Verbal III	Verbal	0.77	0.70
CR	Verbal	1.33	1.15
A	Verbal	0.81	0.74
SC	Verbal	0.76	0.70
Math I	Math	0.56	0.53
Math II	Math	0.56	0.53
Math III	Math	0.41	0.40
Math 4c	Math	0.52	0.50
Math 5c	Math	0.56	0.53
Math S	Math	0.43	0.41
Verbal	Total	4.77	2.41
Math	Total	4.77	2.41

Note: CR = critical reading; A = analogies; SC = sentence completion.

TABLE 7

Proportional Reduction of Mean Squared Error Achieved by Approximation of SAT True Subscores

Subscore	Total Score	$\Psi(R(\tau_X\|\tau_Z)\|L(R(\tau_X\tau_Z)S_X, S_Z)$
Verbal I	Verbal	0.13
Verbal II	Verbal	0.03
Verbal III	Verbal	0.18
CR	Verbal	0.26
A	Verbal	0.16
SC	Verbal	0.14
Math I	Math	0.09
Math II	Math	0.09
Math III	Math	0.08
Math 4c	Math	0.09
Math 5c	Math	0.09
Math S	Math	0.12
Verbal	Total	0.73
Math	Total	0.73

Note: CR = critical reading; A = analogies; SC = sentence completion.

significant figures. The partial regression coefficient $\beta(\tau_X|S_X \cdot S_Z)$ for the critical reading subscore S_X is only 0.17 (see Table 4). Both the linear regression $L(\tau_X|S_Z)$ and the linear regression $L(\tau_X|S_X, S_Z)$ are quite respectable estimates for the true subscore τ_X, for the proportional reductions in MSE are both 0.89 to two significant figures (see Table 5).

In the case of the true residual subscore $R(\tau_X|\tau_Z)$, $\sigma(R(\tau_X|\tau_Z))$, the RMSE for approximation of $R(\tau_X|\tau_Z)$ by its mean of 0, is estimated to be 1.33 (see Table 6), and the RMSE $\sigma(R(R(\tau_X|\tau_Z)|S_X, S_Z))$ from the linear regression of $R(\tau_X|\tau_Z)$ on the observed verbal subscore S_X and the total verbal score S_Z has estimate 1.15, so that there is only a modest gain from approximation of the true residual subscore $R(\tau_X|\tau_Z)$ by the linear regression $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ rather than the trivial constant approximation 0. Note that the proportional reduction in MSE from use of $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ rather than 0 is only 0.26 (see Table 7).

Similar results apply to the other sections of the verbal examination, and similar results also apply if S_X is a section score of the math examination and S_Z is the total score for the math examination. The variations in Table 4 in the partial regression coefficients $\beta(\tau_X|S_X \cdot S_Z)$ mostly just reflect relative lengths of sections. In summary, none of the section scores of SAT I math or SAT I verbal provide any appreciable information concerning an examinee that is not already provided by the math or verbal total score.

On the other hand, the analysis here would certainly support use of separate math and verbal scores. Let S_X be the math total, and let S_Z be the sum of the math and verbal total. As seen in the last row of Table 3, the standard error of measurement $\sigma(e_X) = 3.6$ is the RMSE from use of the observed math score S_X to approximate the true math score τ_X, the RMSE from approximation of the true math total τ_X by Kelley's formula is $\sigma(R(\tau_X|S_X)) = 3.4$, and the RMSE from approximation of the true math total τ_X by linear regression of τ_X on the observed total SAT score S_Z is $\sigma(R(\tau_X|S_Z)) = 5.4$, so that the math true score is much less well predicted by the combined total score than by the math score. There is little value in use of the linear regression $L(\tau_X|S_X, S_Z)$ of the true math score τ_X on both the observed verbal score S_X and the observed SAT total S_Z, for the RMSE $\sigma(R(\tau_X|S_X, S_Z))$ is 3.3, a value only slightly better than was achieved with Kelley's formula.

Similarly, if S_Z remains the total SAT score but S_X is now the SAT verbal score, then, as seen in the second-to-last row in Table 3, the standard error of measurement $\sigma(e_X) = 4.4$ is the MSE from approximation of the true verbal score τ_X by the observed verbal score S_X, the RMSE from approximation of the true verbal score τ_X by Kelley's formula is $\sigma(R(\tau_X|S_X)) = 4.2$, and the RMSE from approximation of τ_X by linear regression of τ_X on the observed total SAT score S_Z is $\sigma(R(\tau_X|S_X, S_Z)) = 5.7$. As in the case of SAT math, there is little value in use of the linear regression $L(\tau_X|S_X, S_Z)$ of the true verbal score τ_X on both the observed verbal score S_X and the observed SAT total S_Z, for the RMSE $\sigma(R(\tau_X|S_X, S_Z))$ is 4.1.

In the case of true residual subscores, for math, as seen in Table 6, $\sigma(R(\tau_X|\tau_Z))$ is 4.77 and $\sigma(R(R(\tau_X|\tau_Z)|S_X, S_Z))$ is 2.41. Because the total score is the sum of the math and verbal subscores, the same results apply for the verbal test. The estimated proportional reduction in MSE from use of the linear regression $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ rather than 0 is 0.73 in both cases (see Table 7). Thus, use of the linear regression does provide a substantial gain over the trivial estimate of 0. The RMSE from use of $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ is about half the corresponding RMSE from use of 0. On the other hand, the proportional reduction of MSE of 0.74 from use of $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ to assess deviation of SAT I math from the value expected by SAT I total is somewhat smaller than the proportional reduction in MSE of 0.92 associated with use of Kelley's formula for estimation of SAT I math.

One may argue that it is unreasonable to expect very much information from subscores in the SAT I math and verbal examinations. The SAT I math and SAT I verbal examinations measure relatively limited content areas. On the other hand, some Praxis examinations contain parts that test very distinct content areas. For instance, consider the test titled *Fundamental Subjects: Content Knowledge* with code 0511 (Grant, 2003). This test measures English language arts (E), mathematics (M), citizenship and social science (C), and science (S). Each area is measured with 25 multiple-choice items, and the total raw score is the sum of the scores for each area. Results are summarized in Tables 8 through 13.

As an approximation to the true subscore τ_X, the linear regression $L(\tau_X|S_X)$ of true subscore τ_X on observed subscore S_X is roughly comparable to the linear regression $L(\tau_X|S_Z)$ of true subscore τ_X on observed total subscore S_Z (see Table 9). For the subscores for citizenship and social science and for science, results from the observed total score are somewhat better than for the results from the observed subscore. The opposite situation applies to subscores for English language arts and mathematics, although the difference for English is too small to be evident in Table 9. Approximation of the true subscore τ_X by the linear regression $L(\tau_X|S_X, S_Z)$ of the true subscore τ_X on both the observed subscore S_X and the observed total score S_Z provides a modest but appreciable improvement in MSE in all cases. Results are best for mathematics, and in all cases, a relatively substantial weight is given to the observed subscore S_X. With the linear regression $L(\tau_X|S_X, S_Z)$ on both observed subscore and observed total score, proportional reductions of MSE are around 0.8 (see Table 11), so that approximations of the true subscores by use of the linear regression $L(\tau_X|S_X, S_Z)$ can be regarded as relatively successful; however, the proportional reductions in error achieved from $L(\tau_X|S_X, S_Z)$ are somewhat smaller than those achieved with $L(\tau_X|S_X, S_Z)$ for any of the subscores of SAT I math or verbal. The essential issue would appear to be that the subscores are less accurately predicted by total score in the Praxis case.

For analysis of the true residual subscore, appreciable but rather modest gains over approximation of the true residual subscore $R(\tau_X|\tau_Z)$ by its mean of 0 are observed in Table 12 for all cases, with best results for English language arts and mathematics. The proportional reductions in MSE reported in Table 13 are all relatively modest.

TABLE 8
Summary Statistics for Praxis Subscores

Subscore	Number of Items	Mean	Standard Deviation	Standard Error of Measurement	Correlation With Total
E	25	21.0	3.3	1.7	.79
M	25	18.2	4.1	1.9	.83
C	25	18.9	3.2	1.8	.81
S	25	15.5	3.6	2.0	.84
Total	100	73.6	11.7	3.7	1.00

Note: E = English language arts; M = mathematics; C = citizenship and social science; S = science.

TABLE 9
Root Mean Squared Error for Approximation of Praxis True Subscores

Subscore	$\sigma(e_X)$	$\sigma(R(\tau_X \vert S_X))$	$\sigma(R(\tau_X \vert S_Z))$	$\sigma(R(\tau_X \vert S_X, S_Z))$
E	1.7	1.5	1.5	1.3
M	1.9	1.7	1.9	1.5
C	1.8	1.5	1.3	1.2
S	2.0	1.7	1.3	1.3

Note: E = English language arts; M = mathematics; C = citizenship and social science; S = science.

TABLE 10
Linear Regression Weights for Approximation for Praxis True Subscores by Observed Subscores and Observed Total Scores

Subscore	$\beta(\tau_X \vert S_X \cdot S_Z)$	$\beta(\tau_X \vert S_Z \cdot S_X)$
E	0.44	0.10
M	0.51	0.12
C	0.28	0.14
S	0.22	0.17

Note: E = English language arts; M = mathematics; C = citizenship and social science; S = science.

TABLE 11
Proportional Reduction of Mean Squared Error Achieved by Approximation of Praxis True Subscores

Subscore	$\Psi(\tau_X \vert L(\tau_X \vert S_X))$	$\Psi(\tau_X \vert L(\tau_X \vert S_Z))$	$\Psi(\tau_X \vert L(\tau_X \vert S_X, S_Z))$
E	0.73	0.70	0.80
M	0.79	0.73	0.83
C	0.68	0.77	0.81
S	0.69	0.80	0.82

Note: E = English language arts; M = mathematics; C = citizenship and social science; S = science.

TABLE 12
Root Mean Squared Error for Approximation of Praxis True Residual Subscores

Subscore	$\sigma(R(\tau_X \mid \tau_Z))$	$\sigma(R(R(\tau_X \mid \tau_Z) \mid S_X, S_Z))$
E	1.32	1.00
M	1.58	1.13
C	1.01	0.85
S	0.98	0.85

Note: E = English language arts; M = mathematics; C = citizenship and social science; S = science.

TABLE 13
Proportional Reduction of Mean Squared Error Achieved by Approximation of Praxis True Residual Subscores

Subscore	$\Psi(R(\tau_X \mid \tau_Z) \mid L(R(\tau_X \mid \tau_Z) \mid S_X, S_Z))$
E	0.43
M	0.48
C	0.29
S	0.25

Note: E = English language arts; M = mathematics; C = citizenship and social science; S = science.

4. Conclusions

The methods of subscore analysis proposed are very easily implemented and provide a rational criterion for assessing the value of subscores. Results suggest that a good deal of caution is needed. With the criterion of MSE of approximation of true subscores, observed subscores are most likely to have value if they have relatively high reliability by themselves and if the true subscore and true total score have only a moderate correlation. Both conditions are important. The SAT subscores are relatively unsuccessful due to the very high correlations of their true scores with the true total score; however, many of the subscores are rather reliable, so that these observed scores can be employed by themselves to provide relatively accurate approximations to the true subscores. Nonetheless, a linear regression of true subscore on both observed subscore and observed total score gives very high weight to the total score and does lead to an appreciably better approximation of true subscore than is provided by observed subscore alone.

The Praxis subscores are often less reliable than are many of the SAT subscores, but the correlation of true subscores to true total score is somewhat more modest than for the SAT subscores. Nonetheless, even for the Praxis subscores, which are all based on 25 items and measure very different content areas, the subscores are best used when combined with the total score, and the proportional reduction in MSE from use of the linear regression $L(\tau_X \mid S_X, S_Z)$ to predict the true subscore is somewhat less than is the proportional reduction in MSE from use of Kelley's formula to approximate the true total score from the observed total score.

Although the results here do not prove that subscores cannot be useful, they do suggest that claims for the value of subscores should be treated skeptically and should be verified by use of procedures similar to those in this report. The results also suggest that a linear combination of observed subscore and observed total score may at times be somewhat more useful than an observed subscore.

This report emphasizes simple approaches to subscores. It is possible that alternatives can be constructed that are quite attractive in particular applications. For example, subscore predictions from total scores may be based on use of log-linear models or use of item-response theory. Thus, additional work can be considered to aid in subscore assessment.

Appendix
Derivations of Formulas

Linear Regression of True Subscore on Observed Subscore and Observed Total Score

To verify that $|\rho(S_X, S_Z)| < 1$, consider real constants a and b. The variance $\sigma^2(aS_X + bS_Z)$ of a linear combination $aS_X + bS_Z = (a+b)S_X + bS_Y$ is

$$(a+b)^2\sigma^2(\tau_X) + b^2\sigma^2(\tau_Y) + 2(a+b)b\sigma(\tau_X)\sigma(\tau_Y)\rho(\tau_X, \tau_Y) + (a+b)^2\sigma(e_X) + b^2\sigma^2(e_Y)$$

(Lord & Novick, 1968), so that

$$\sigma^2(aS_X + bS_Z) = a^2\sigma^2(S_X) + 2ab\sigma(S_X)\sigma(S_Z)\rho(S_X, S_Z) + b^2\sigma(S_Z)$$

is only 0 if $a = b = 0$. The case $a = 1$ and $b = -\beta(S_X|S_Z) = -\rho(S_X, S_Z)\sigma(S_X)/\sigma(S_Z)$ yields

$$\sigma^2(aS_X + bS_Z) = \sigma^2(R(S_X|S_Z)) = \sigma^2(S_X)[1 - \rho^2(S_X, S_Z)] > 0,$$

so that $|\rho(S_X, S_Z)| < 1$.

Linear Regression of True Residual Subscore on Observed Subscore and Observed Total Score

The linear regression $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ of $R(\tau_X|\tau_Z)|S_X, S_Z)$ can be found by an argument very similar to that in Section 1.3. One has partial regression coefficients

$$\beta(R(\tau_X|\tau_Z)|S_X \cdot S_Z) = \frac{\sigma(R(\tau_X\tau_Z))[\rho(S_X, R(\tau_X|\tau_Z)) - \rho(R(\tau_X|\tau_Z), S_Z)\rho(S_X, S_Z)]}{\sigma(S_X)[1 - \rho^2(S_X, S_Z)]}$$

and

$$\beta(R(\tau_X|\tau_Z)|S_Z \cdot S_X) = \frac{\sigma(R(\tau_X|\tau_Z))[\rho(R(\tau_X|\tau_Z), S_Z) - \rho(S_X, R(\tau_X|\tau_Z))\rho(S_X, S_Z)]}{\sigma(S_Z)[1 - \rho^2(S_X, S_Z)]}.$$

Because the errors of measurement e_X and e_Z are uncorrelated with the true scores τ_X and τ_Z, the errors e_X and e_Z are also uncorrelated with

$$R(\tau_X|\tau_Z) = [\tau_X - E(\tau_X)] - \beta(\tau_X|\tau_Z)[\tau_Z - E(\tau_Z)]$$

(Lord & Novick, 1968). Standard attenuation arguments then imply that

$$\rho(S_X, R(\tau_X|\tau_Z)) = \rho(S_X, \tau_X)\rho(\tau_X, R(\tau_X|\tau_Z))$$

and

$$\rho(S_Z, R(\tau_X|\tau_Z)) = \rho(S_Z, \tau_X)\rho(\tau_Z, R(\tau_X|\tau_Z))$$

(Lord & Novick, 1968). The true residual subscore $R(\tau_X|\tau_Z)$ is uncorrelated with the true total score τ_Z (Lord & Novick, 1968), so that $\rho(S_Z, R(\tau_X|\tau_Z))$ and $\rho(\tau_Z, R(\tau_X|\tau_Z))$ are both 0. The covariance $\mathrm{Cov}(\tau_X, R(\tau_X|\tau_Z))$ of the true subscore τ_X and the true residual subscore $R(\tau_X|\tau_Z))$ is the MSE $\sigma^2(R(\tau_X|\tau_Z))$ (Cramér, 1946), so that the correlation $\rho(\tau_X, R(\tau_X|\tau_Z))$ of τ_X and $R(\tau_X|\tau_Z)$ is $\sigma(R(\tau_X|\tau_Z))/\sigma(\tau_X)$. Recall that $\sigma(\tau_X) = \rho(S_X, \tau_X)\sigma(S_X)$, $\sigma^2(R(S_X|S_Z)) = \sigma^2(S_X)[1 - \rho^2(S_X, S_Z)]$, and $\beta(S_X|S_Z) = \sigma(S_X)\rho(S_X, S_Z)/\sigma(S_Z)$. Thus,

$$\begin{aligned}\beta(R(\tau_X|\tau_Z)|S_X \cdot S_Z) &= \frac{\sigma^2(R(\tau_X|\tau_Z))\rho(S_X, \tau_X)}{\sigma(\tau_X)\sigma(S_X)[1 - \rho^2(S_X, S_Z)]} \\ &= \frac{\sigma^2(R(\tau_X|\tau_Z))}{\sigma^2(R(S_X|S_Z))}\end{aligned}$$

and

$$\begin{aligned}\beta(R(\tau_X|\tau_Z)|S_Z \cdot S_X) &= -\frac{\sigma^2(R(\tau_X|\tau_Z))\rho(S_X, \tau_X)\rho(S_X, S_Z)}{\sigma(\tau_X)\sigma(S_Z)[1 - \rho^2(S_X, S_Z)]} \\ &= -\frac{\sigma(R(\tau_X|\tau_Z))\beta(S_X|S_Z)}{\sigma^2(R(S_X|S_Z))},\end{aligned}$$

so that

$$L(R(\tau_X|\tau_Z)|S_X, S_Z) = \frac{\sigma^2(R(\tau_X|\tau_Z))}{\sigma^2(R(S_X|S_Z))} R(S_X|S_Z).$$

Because $L(R(\tau_X|\tau_Z)|S_X, S_Z)$ is a linear function of the observed residual subscore $R(S_X|S_Z)$, the MSE $\sigma^2(R(R(\tau_X|\tau_Z)|S_X, S_Z))$ from the linear regression of the true residual subscore $R(\tau_X|\tau_Z)$ on the observed subscore S_X and the observed total score S_Z is at least as great as the MSE $\sigma^2(R(R(\tau_X|\tau_Z)|R(S_X|S_Z))$ from the regression of $R(\tau_X|\tau_Z)$ on the observed residual subscore $R(S_X|S_Z)$. Because $R(S_X|S_Z)$ is a linear combination $[S_X - E(S_X)] - \beta(S_X|S_Z)[S_Z - E(S_Z)]$ of S_X and S_Z and $L(R(\tau_X|\tau_Z)|R(S_X|S_Z)) = \beta(R(\tau_X|\tau_Z)|R(S_X|S_Z))R(S_X|S_Z)$, the MSE $\sigma^2(R(R(\tau_X|\tau_Z)|S_X, S_Z))$ is less than the MSE $\sigma^2(R(R(\tau_X|\tau_Z)|R(S_X|S_Z))$ from the regression of $R(\tau_X|\tau_Z)$ on the observed residual subscore $R(S_X|S_Z)$ unless $\beta(R(\tau_X|\tau_Z)|S_X \cdot S_Z) = \beta(R(\tau_X|\tau_Z)|R(S_X|S_Z))$ and $\beta(R(\tau_X|\tau_Z)|S_Z \cdot S_X) = -\beta(\tau_X|\tau_Z)\beta(R(\tau_X|\tau_Z)|R(S_X|S_Z))$. It follows that

$$L(R(\tau_X|\tau_Z)|S_X, S_Z) = L(R(\tau_X|\tau_Z)|R(S_X|S_Z))$$

and

$$\beta(R(\tau_X|\tau_Z)|R(S_X|S_Z)) = \frac{\sigma^2(R(\tau_X|\tau_Z))}{\sigma^2(R(S_X|S_Z))}.$$

The correlation coefficient $\rho(R(\tau_X|\tau_Z), R(S_X|S_Z))$ is then $\sigma(R(\tau_X|\tau_Z))/\sigma(R(S_X|S_Z))$, so that the MSE

$$\sigma^2(R(\tau_X|\tau_Z)|S_X, S_Z) = \sigma^2(R(\tau_X|\tau_Z))[1 - \Psi(R(\tau_X|\tau_Z)|S_X, S_Z)],$$

where

$$\Psi(R(\tau_X|\tau_Z)|S_X, S_Z) = \frac{\sigma^2(R(\tau_X|\tau_Z))}{\sigma^2(R(S_X|S_Z))}$$

must be between 0 and 1. If $\Psi(R(\tau_X|\tau_Z)|S_X, S_Z)$ is 1, then $\sigma^2(R(R(\tau_X|\tau_Z)|R(S_X|S_Z)) = 0$ and $R(S_X|S_Z) = R(\tau_X|\tau_Z)$ with probability 1. Because

$$R(S_X|S_Z) = [\tau_X - E(\tau_X)] - \beta(S_X|S_Z)[\tau_Z - E(\tau_Z)] + e_X - \beta(S_X|S_Z)e_Z,$$

where the errors of measurement e_X and e_Z are uncorrelated with the true scores τ_X and τ_Z and $e_X - \beta(S_X|S_Z)e_Z = [1 - \beta(S_X|S_Z)]e_X - \beta(S_X|S_Z)e_Y$ has positive variance $\sigma^2(e_X)[1 - \beta(S_X|S_Z)]^2 + \sigma^2(e_Y)[\beta(S_X|S_Z)]^2$, it is not possible that $R(S_X|S_Z) = R(\tau_X|\tau_Z)$ with probability 1. Thus, $\Psi(R(\tau_X|\tau_Z)|S_X, S_Z) < 1$.

Estimation of the Correlation of the True Subscore and the True Total Score

The true subscore τ_X and the true total score τ_Z have correlation

$$\rho(\tau_X, \tau_Z) = \frac{\mathrm{Cov}(\tau_X, \tau_Z)}{\sigma(\tau_X)\sigma(\tau_Z)}.$$

The standard deviation $\sigma(\tau_X)$ of the true subscore τ_X may be estimated by the square root $\hat{\sigma}(\tau_X)$ of the estimated variance $\hat{\sigma}^2(\tau_X)$ of τ_X, whereas the standard deviation $\sigma(\tau_Z)$ of the true total score τ_Z may be estimated by the square root $\hat{\sigma}(\tau_Z)$ of the estimated variance $\hat{\sigma}^2(\tau_Z)$ of τ_Z. Because the true total score τ_Z is the sum of the true subscore τ_X and the true remainder score τ_Y, the covariance $\mathrm{Cov}(\tau_X, \tau_Z)$ of the true subscore τ_X and the true total score τ_Z satisfies

$$\mathrm{Cov}(\tau_X, \tau_Z) = \sigma^2(\tau_X) + \mathrm{Cov}(\tau_X, \tau_Y).$$

Similarly, because the observed total score S_Z is the sum of the observed subscore S_X and the observed remainder score S_Y, the covariance $\mathrm{Cov}(S_X, S_Z)$ of the observed subscore S_X and the observed total score S_Z satisfies

$$\mathrm{Cov}(S_X, S_Z) = \sigma^2(S_X) + \mathrm{Cov}(S_X, S_Y).$$

Because the error of measurement e_X for the subscore S_X and the error of measurement e_Y for the remainder score S_Y are uncorrelated, standard attenuation formulas imply that the covariance $\mathrm{Cov}(\tau_X, \tau_Y)$ of the true subscore τ_X and the true remainder score τ_Y is the same as the covariance $\mathrm{Cov}(S_X, S_Y)$ of the observed subscore S_X and the observed remainder score S_Y (Holland & Hoskens, 2003). Recall that $\sigma^2(S_X) = \sigma^2(\tau_X) + \sigma^2(e_X)$. It follows that

$$\mathrm{Cov}(\tau_X, \tau_Z) = \mathrm{Cov}(S_X, S_Z) - \sigma^2(e_X).$$

Thus,

$$\rho(\tau_X, \tau_Z) = \frac{\mathrm{Cov}(S_X, S_Z) - \sigma^2(e_X)}{\sigma(\tau_X)\sigma(\tau_Z)}$$

may be estimated by

$$\begin{aligned}\hat{\rho}(\tau_X, \tau_Z) &= \frac{\hat{\sigma}(S_X)\hat{\sigma}(S_Z)\hat{\rho}(S_X, S_Z) - \hat{\sigma}^2(e_X)}{\hat{\sigma}(\tau_X)\hat{\sigma}(\tau_Z)} \\ &= \frac{\hat{\rho}(S_X, S_Z)}{\hat{\rho}(S_X, \tau_X)\hat{\rho}(S_Z, \tau_Z)} - \frac{\hat{\sigma}^2(e_X)}{\hat{\sigma}(\tau_X)\hat{\sigma}(\tau_Z)}.\end{aligned}$$

References

Cramér, H. (1946). *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press.

Dressel, P. L. (1940). Some remarks on the Kuder-Richardson 20 (KR-20) reliability statistic for formula scored tests. *Psychometrika*, *5*, 305–310.

Feigenbaum, M., & Hammond, S. (2003). *Test analysis, College Board, SAT I: Reasoning Test, fall 2002 administrations, 3YSA03–3YSA05* (Report No. SR-2003–37). Princeton, NJ: Educational Testing Service.

Grant, M. (2003). *Fundamental subjects: Content knowledge (0511), test analysis, form 3ypx1* (Report No. SR-2003–62). Princeton, NJ: Educational Testing Service.

Holland, P. W., & Hoskens, M. (2003). Classical test theory as a first-order item response theory: Application to true-score prediction from a possibly nonparallel test. *Psychometrika*, *68*, 123-149.

Kelley, T. L. (1947). *Fundamentals of statistics*. Cambridge, MA: Harvard University Press.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*, 151–160.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Wainer, H., Vevea, J. L., Camacho, F., Reeve, B. B., Swygert, K. A., & Thissen, D. (2001). Augmented scores—"Borrowing strength" to compute scores based on small numbers of items. In D. Thissen & H. Wainer (Eds.), *Test scoring* (pp. 343–387). Mahwah, NJ: Lawrence Erlbaum.

Yen, W. M. (1987, June). *A Bayesian/IRT index of objective performance*. Paper presented at the annual meeting of the Psychometric Society, Montreal, Quebec.

Author

SHELBY J. HABERMAN is the research director, Statistical and Psychometric Theory and Practice, Mailstop 12-T, Educational Testing Service, Rosedale Road, Princeton, NJ 08541; SHaberman@ets.org. His principal research interests are analysis of qualitative data, asymptotic approximations, and application of statistics to educational measurement.

Manuscript received March 30, 2006
Revision received December 12, 2006
Accepted March 4, 2007