Advances in Theoretical and Applied Mathematics.
ISSN 0973-4554 Volume 21, Number 2 (2026) pp. 139-155
©Research India Publications
https://www.ripublication.com

A Mathematical Model of Evaluator-Related Variation in Continuous Internal Assessment

Sudipto Roy[1*]

[1]*Department of Physics, St. Xavier's College, Kolkata,, West Bengal, India*
e-mail: roy.sudipto@sxccal.edu
[*]*ORCID: https://orcid.org/0000-0002-8811-2511*

Abstract

Continuous Internal Assessment (CIA) is widely practised in colleges and universities, but students of similar merit, ability, and diligence may receive different marks because of differences in teachers' marking tendencies. This study presents a model for a paper with three components, evaluated by three teachers. One teacher is taken as a reference for comparison, and the marking tendencies of the other two teachers are represented by parameters. The paper compares three arrangements: one teacher evaluates all three components, two teachers share the components in the sense that one teacher evaluates two components and another evaluates the remaining one, and three different teachers evaluate one component each. For each of these three arrangements, the variation in the possible total marks of a given student is determined. An analytical calculation shows that this variation is the smallest when three different teachers evaluate the three components, the largest when one teacher evaluates all three components, and intermediate when two teachers share the components. The paper also shows how the parameters representing the marking tendencies of evaluators can be estimated from the marks awarded. The model provides a framework for understanding evaluator-related variation in internal assessment and for studying how shared evaluation may reduce the influence of individual marking styles. Further studies can be carried out using this framework and real assessment data.

Keywords: Continuous Internal Assessment, Generalizability Theory, Evaluator-Related Variation, Marking Tendencies, Assessment Reliability, Mathematical Modelling.

Mathematics Subject Classification (2020): 97C70, 97D60 (primary); 62P25 (secondary).

1. INTRODUCTION

Assessment is an important part of teaching and learning. It helps teachers to understand the progress of students, and it shows how well students have understood what has been taught [1].

Continuous Internal Assessment (CIA) has now become an important part of the evaluation system at the college and university levels. It is used to evaluate students during a course. It may include class tests, viva voce, assignments, presentations, projects, laboratory work, or other academic activities. But different teachers may judge similar work in different ways. As a result, the same student can receive different marks from different teachers. One teacher may be relatively strict, while another may be relatively lenient. Such differences can affect the final assessment.

Differences among evaluators are a well-known concern in educational assessment. Research on assessment reliability shows that an evaluator can be a source of variation in scores [2]. One established framework for studying such sources of variation is Generalizability Theory [3]. It examines how much different factors (students, evaluators, components of an assessment, and their interactions) contribute to the overall variation in scores. The distinction between reliability and validity is also important when assessment instruments are examined [4]. Studies have applied Generalizability Theory to the reliability of ratings and essay assessment [2, 5].

The present paper addresses the related problem of evaluator-dependent variation in marks through a simple mathematical model. It is not a formal application of Generalizability Theory: it does not estimate variance components and it does not calculate a generalizability coefficient. Readers who are unfamiliar with Generalizability Theory can treat the model that follows as a self-contained mathematical description; no specialised knowledge of that theory is required to follow the rest of the paper.

This paper develops a simple mathematical model with three teachers and three components of a paper. One teacher is taken as a reference teacher, and the marking tendencies of the other two teachers are represented by parameters. The paper compares three arrangements: one teacher evaluates all three components; two teachers share the components, in the sense that one teacher evaluates two components and another evaluates the remaining one; and three different teachers evaluate one component each. An analytical calculation compares the variation in the total marks for these three arrangements, under the assumptions of the model. The paper also presents a simple method for estimating the marking parameters from the marks given by the teachers.

The main aim is to compare the variation in the total marks for a given student for the three evaluation arrangements, within the proposed model. The second aim is to show how the parameters representing the marking tendencies of teachers can be estimated from the marks given by the teachers.

2. MODEL FORMULATION

2.1. Symbols Used

I) Identities of three teachers: T_1, T_2, T_3.

II) Three CIA components of the paper: C_1, C_2, C_3.

III) Full marks for the three components of the paper: F_1, F_2, F_3.

IV) Marks given by T_x for C_y are denoted by M_{xy}, where $x, y = 1, 2, 3$.

V) The parameters a and b are introduced through the expressions in Table 1.

2.2. Expressions for Marks

Table 1 lists the marks given by T_1, T_2, and T_3 to a student for the three components of the paper, C_1, C_2, and C_3. Marks given by T_2 and T_3 are expressed in terms of marks given by T_1.

Table 1: Expressions for marks given by the teachers for three components

Teacher	C_1	C_2	C_3
T_1	M_{11}	M_{12}	M_{13}
T_2	$M_{21} = (1-a)M_{11}$	$M_{22} = (1-a)M_{12}$	$M_{23} = (1-a)M_{13}$
T_3	$M_{31} = M_{11}$ $+b(F_1 - M_{11})$	$M_{32} = M_{12}$ $+b(F_2 - M_{12})$	$M_{33} = M_{13}$ $+b(F_3 - M_{13})$

Note: Here $0 < a, b \leq 1$. The cases $a = 0$ and $b = 0$ are excluded by the assumptions in Section 2.3. The strict bound $M_{1i} < F_i$ ensures that T_3 gives a mark strictly higher than that given by T_1, while the expression for T_3 never exceeds the full marks of the corresponding component. Smaller values of a and b indicate that the marking tendencies of T_2 and T_3, respectively, are closer to that of the reference teacher T_1.

Here, a represents the relative reduction in marks by T_2, while b represents the relative increase in marks towards the full marks by T_3. Thus, a and b measure the marking tendencies of the two non-reference teachers, relative to T_1.

2.3. Assumptions

I) T_1 gives exactly the marks a student deserves.

II) T_2 gives lower marks than T_1.

III) T_3 gives higher marks than T_1.

IV) For each component, $0 < M_{1i} < F_i$, where $i = 1, 2, 3$, and the parameters satisfy $0 < a, b \leq 1$.

V) For a given student, the parameters a and b are taken to be the same for all three components.

The first assumption gives T_1 the role of an ideal reference teacher. However, the mathematical analysis uses the marks given by T_1 only as a reference for comparison. It does not require independent proof that these marks are objectively the marks that a student deserves.

3. MATHEMATICAL ANALYSIS AND NUMERICAL CALCULATIONS

3.1. Calculation of marks based on a and b

Table 2 lists the values chosen for all the quantities in this model. For the first calculations shown in Tables 3 and 4, we have chosen $a = b = 0.2$.

Table 2: Values chosen for calculations shown in Tables 3 and 4

Components	**Full Marks**	**Marks Given by** T_1
C_1	$F_1 = 10$	$M_{11} = 7.0$
C_2	$F_2 = 15$	$M_{12} = 10.0$
C_3	$F_3 = 25$	$M_{13} = 18.0$
Chosen parameters: $a = b = 0.2$		

Table 3 lists the total marks obtained by the student for the three possible cases in which one teacher evaluates all three components. Marks have been calculated using Tables 1 and 2.

Table 3: Total marks when one teacher evaluates all three components

Teacher for C_1	**Teacher for** C_2	**Teacher for** C_3	**Marks**
T_1	T_1	T_1	35.0
T_2	T_2	T_2	28.0
T_3	T_3	T_3	38.0
		Mean	**33.67**
		Standard Deviation (S)	**4.19**

Table 4 lists the total marks obtained by the student for the six possible assignments in which three different teachers evaluate the three components. Each teacher evaluates one component. Marks have been calculated using Tables 1 and 2.

Here we consider one student and compare the possible teacher assignments. For each evaluation arrangement, its cases are considered separately, and all the cases within that arrangement are given equal weight. Thus, the three one-teacher cases are used to calculate S, the six three-teacher cases are used to calculate S', and the 18 intermediate cases are used to calculate S''. Equal weighting means that every possible assignment is treated as equally important in the calculation, irrespective of how frequently that assignment might occur in an actual assessment system. The cases belonging to different arrangements are not mixed when we calculate these standard deviations. Accordingly, the standard deviations are population standard deviations over the complete set of assignments for each arrangement. They measure the variation in the total caused by the teacher assignment, not the variation between different students. They are not sample standard deviations.

Table 4: Total marks when three teachers evaluate the three components

Teacher for C_1	**Teacher for C_2**	**Teacher for C_3**	**Marks**
T_1	T_2	T_3	34.4
T_1	T_3	T_2	32.4
T_2	T_1	T_3	35.0
T_2	T_3	T_1	34.6
T_3	T_1	T_2	32.0
T_3	T_2	T_1	33.6
		Mean	**33.67**
		Standard Deviation (S')	**1.12**

If N denotes the number of possible assignments in the evaluation arrangement under consideration, the population standard deviation for one-teacher evaluation is calculated by the following formula:

$$S = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2}. \tag{1}$$

The quantity under the square root is the population variance: the average of the squared deviations of the totals from their mean. The square root brings this average squared deviation back to the same units as the marks themselves, and it gives the population standard deviation (also called the root-mean-square deviation). It therefore measures, in a single number, how widely the totals are spread around their mean.

For three-teacher evaluation, the corresponding population standard deviation is calculated by the following formula:

$$S' = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(Y_k - \bar{Y})^2}. \tag{2}$$

For one-teacher evaluation, $N = 3$; for three-teacher evaluation, $N = 6$; and for the intermediate arrangement, $N = 18$. Here, X_i denotes the total mark in the ith one-teacher case, Y_k denotes the total mark in the kth three-teacher case, and Z_i denotes the total mark in the ith intermediate case. The index k labels the six assignments and is distinct from the component indices i and j used in Y_{ij} below. The symbols S, S', and S'' denote, respectively, the standard deviations for one-teacher evaluation, three-teacher evaluation, and the intermediate arrangement.

Inference based on Tables 3 and 4: The standard deviation is 4.19 when one teacher evaluates all three components and 1.12 when three different teachers evaluate the three components of the paper.

3.1.1. Analytical comparison of S and S'

The numerical comparison in Tables 3 and 4 can be made stronger by deriving S and S' analytically for the same student. The derivation below uses the assumptions stated in Section 2.3. Let us define the amount by which T_2 lowers the mark for component i, and the amount by which T_3 raises it:

$$A_i = aM_{1i}, \qquad B_i = b(F_i - M_{1i}), \qquad i = 1, 2, 3. \tag{3}$$

Thus, A_i is the reduction caused by T_2, while B_i is the increase caused by T_3.
Let the reference total and the sums of these quantities be

$$R = \sum_{i=1}^{3} M_{1i}, \qquad A = \sum_{i=1}^{3} A_i, \qquad B = \sum_{i=1}^{3} B_i. \tag{4}$$

For one-teacher evaluation, the three possible total marks are

$$X_1 = R, \qquad X_2 = R - A, \qquad X_3 = R + B. \tag{5}$$

Their mean is

$$\bar{X} = R + \frac{B - A}{3}. \tag{6}$$

Therefore, the deviations of the three cases from their mean are

$$X_1 - \bar{X} = \frac{A - B}{3}, \qquad X_2 - \bar{X} = -\frac{2A + B}{3}, \qquad X_3 - \bar{X} = \frac{A + 2B}{3}. \tag{7}$$

Using the population-variance formula corresponding to the standard deviation defined earlier, the variance for one-teacher evaluation is

$$S^2 = \frac{1}{3}\left[\left(\frac{A - B}{3}\right)^2 + \left(\frac{2A + B}{3}\right)^2 + \left(\frac{A + 2B}{3}\right)^2\right] = \frac{2(A^2 + AB + B^2)}{9}. \tag{8}$$

Thus,

$$S = \sqrt{\frac{2(A^2 + AB + B^2)}{9}}. \tag{9}$$

For three-teacher evaluation, let Y_{ij} denote the total mark when T_2 evaluates component i, T_3 evaluates component j, and T_1 evaluates the remaining component, where $i \neq j$. Then

$$Y_{ij} = R - A_i + B_j, \qquad i \neq j. \tag{10}$$

There are six such assignments, in which each of the three teachers evaluates exactly one component. Since each A_i and each B_i occurs exactly twice among these six cases, their mean is the same as for the one-teacher cases:

$$\bar{Y} = R + \frac{B - A}{3}. \tag{11}$$

Substituting the six values of Y_{ij} into the population-variance formula gives

$$S'^2 = \frac{1}{6} \sum_{\substack{i,j=1 \\ i \neq j}}^{3} (Y_{ij} - \bar{Y})^2. \tag{12}$$

Since $Y_{ij} = R - A_i + B_j$ and $\bar{Y} = R + (B - A)/3$, we have

$$Y_{ij} - \bar{Y} = -A_i + B_j - \frac{B - A}{3}.$$

In the six assignments, each A_i^2 and each B_i^2 occurs twice, while each cross-product $A_i B_j$ with $i \neq j$ occurs in exactly one assignment. Using the identity

$$\frac{1}{6} \sum (Y_{ij} - \bar{Y})^2 = \frac{1}{6} \sum Y_{ij}^2 - \bar{Y}^2,$$

and substituting the expressions for Y_{ij} and $\bar{Y}$, we obtain

$$6S'^2 = 2 \sum_{i=1}^{3} A_i^2 + 2 \sum_{i=1}^{3} B_i^2 - 2 \sum_{\substack{i,j=1 \\ i \neq j}}^{3} A_i B_j - \frac{2}{3}(B - A)^2.$$

For the comparison of the two variances, let us define the three sums of cross-products that occur when A^2, B^2, and AB are expanded. These are the cross-terms obtained when the sums of the component-wise quantities are squared or multiplied:

$$\begin{aligned} P_A &= A_1 A_2 + A_1 A_3 + A_2 A_3, \\ P_B &= B_1 B_2 + B_1 B_3 + B_2 B_3, \\ P_{AB} &= A_1 B_2 + A_1 B_3 + A_2 B_1 + A_2 B_3 + A_3 B_1 + A_3 B_2. \end{aligned} \tag{13}$$

The quantities P_A, P_B, and P_{AB} collect the cross-products between different components. We use them to expand A^2, B^2, and AB below.

To simplify this expression, we first write the squares and product of the sums A and B in terms of the individual component quantities. From the definitions in Eq. (4),

$$\begin{aligned} A^2 &= \sum_{i=1}^{3} A_i^2 + 2P_A, \\ B^2 &= \sum_{i=1}^{3} B_i^2 + 2P_B, \\ AB &= \sum_{i=1}^{3} A_i B_i + P_{AB}. \end{aligned} \tag{14}$$

Also, the sum of the cross-products with unequal component indices is

$$\sum_{\substack{i,j=1 \\ i \neq j}}^{3} A_i B_j = P_{AB}.$$

Substituting these expressions into the preceding equation and collecting the terms gives

$$6S'^2 = \frac{4}{3}\left(\sum_{i=1}^{3} A_i^2 + \sum_{i=1}^{3} B_i^2 + \sum_{i=1}^{3} A_i B_i - P_A - P_B\right) - \frac{2}{3}P_{AB}.$$

Using Eq. (14) to replace the cross-product sums, we obtain

$$S'^2 = \frac{3\left(\sum_{i=1}^{3} A_i^2 + \sum_{i=1}^{3} B_i^2 + \sum_{i=1}^{3} A_i B_i\right) - (A^2 + AB + B^2)}{9}. \tag{15}$$

Using the expansions in Eq. (14), together with Eqs. (8) and (15), the difference between the two variances is

$$S^2 - S'^2 = \frac{2P_A + P_{AB} + 2P_B}{3}. \tag{16}$$

Assumption IV gives $0 < M_{1i} < F_i$ directly. Since $a > 0$ and $b > 0$, it follows that $A_i = aM_{1i} > 0$ and $B_i = b(F_i - M_{1i}) > 0$ for each $i = 1, 2, 3$. Hence, P_A, P_B, and P_{AB} are all positive. Therefore, $2P_A + P_{AB} + 2P_B > 0$, and

$$S^2 - S'^2 > 0. \tag{17}$$

Since S and S' are standard deviations, they are non-negative. Eq. (17) therefore gives $S' < S$ under the stated conditions. Thus, the reduction in variation follows from the model itself. It is not merely a numerical feature of the examples in Table 6.

3.1.2. Analysis of the Intermediate Evaluation Arrangement

We now consider a third arrangement. In this arrangement, one teacher evaluates two components and another teacher evaluates the remaining component. The third teacher does not evaluate any component in that particular assignment. There are 18 possible assignments of this type. To count them, we first choose the teacher who evaluates two components (3 choices), then choose the component evaluated by the other teacher (3 choices), and finally choose that other teacher (2 choices). Thus, the total number of assignments is $3 \times 3 \times 2 = 18$. We give equal weight to all these assignments, as we did for the six assignments considered above.

Let Z denote the total mark in any one of these 18 assignments. We use the same quantities A_i and B_i defined in Eq. (3). For a component evaluated by T_1, the change from the reference mark is zero; for a component evaluated by T_2, the change is $-A_i$; and for a component evaluated by T_3, the change is B_i.

To see the 18 cases clearly, suppose that component j is evaluated by the teacher who evaluates only one component. Let p and q denote the other two components, so that $p \neq q$, $p \neq j$, and $q \neq j$. The six possible changes in the total mark, relative to R, are

$$\begin{aligned} &-A_j, \quad B_j, \quad -(A_p + A_q), \\ &B_p + B_q, \quad -(A_p + A_q) + B_j, \\ &B_p + B_q - A_j. \end{aligned} \tag{18}$$

These six changes represent all possible ways of combining a pair of components with the remaining single component, for a fixed choice of the single component. These six cases correspond, respectively, to the following choices: the pair is evaluated by T_1 and the single component by T_2; the pair is evaluated by T_1 and the single component by T_3; the pair is evaluated by T_2 and the single component by T_1; the pair is evaluated by T_3 and the single component by T_1; the pair is evaluated by T_2 and the single component by T_3; and the pair is evaluated by T_3 and the single component by T_2. The index j can be 1, 2, or 3. Hence, Eq. (18) describes all 18 assignments.

In these 18 cases, each component is evaluated by each teacher the same number of times. As a result, for any fixed component i, the reference mark M_{1i} appears together with a reduction A_i in six of the assignments, together with an increase B_i in six other assignments, and with neither change (the reference teacher evaluates the component) in the remaining six assignments. The average contribution of component i is therefore

$$M_{1i} + \frac{6(-A_i) + 6(B_i) + 6(0)}{18} = M_{1i} + \frac{B_i - A_i}{3}.$$

Adding the average contributions of the three components yields the mean total

$$\bar{Z} = R + \frac{B - A}{3}. \tag{19}$$

This is the same mean as in the two arrangements already considered.

The population variance is obtained by taking the squared difference of each of the 18 totals from this mean, adding the 18 squared differences, and dividing by 18:

$$S''^2 = \frac{1}{18} \sum_{\text{18 assignments}} (Z - \bar{Z})^2. \tag{20}$$

For a fixed choice $j = 1$, with $p = 2$ and $q = 3$, the six deviations are

$$-A_1 - \frac{B - A}{3}, \quad B_1 - \frac{B - A}{3}, \quad -(A_2 + A_3) - \frac{B - A}{3},$$

$$B_2 + B_3 - \frac{B - A}{3}, \quad -(A_2 + A_3) + B_1 - \frac{B - A}{3}, \quad B_2 + B_3 - A_1 - \frac{B - A}{3}.$$

The corresponding six deviations for $j = 2$ and $j = 3$ are obtained by changing the component indices. For each assignment, let Δ denote the corresponding change in

the total mark relative to R. Adding all 18 squared deviations and collecting the terms belonging to the same component gives

$$\sum_{j=1}^{3} \sum_{\text{six cases for } j} \left(\Delta - \frac{B-A}{3}\right)^2 = 4\sum_{i=1}^{3} \left(A_i^2 + A_i B_i + B_i^2\right), \tag{21}$$

In other words, the left-hand side is the sum of the 18 squared differences from the mean. To see the cancellation, collect the coefficients of the products involving different components after adding the three groups corresponding to $j = 1, 2, 3$. Each pair of distinct components enters these three groups symmetrically, so the cross-component contributions cancel when all three groups are added. The coefficient of each such product is therefore zero. What remains are only the component-wise terms A_i^2, $A_i B_i$ and B_i^2; each of these terms appears exactly four times, which accounts for the overall factor of 4. Dividing this result by 18, as required by Eq. (20), gives

$$S''^2 = \frac{2}{9}\sum_{i=1}^{3} \left(A_i^2 + A_i B_i + B_i^2\right). \tag{22}$$

Thus, the variance for this intermediate arrangement depends only on the sum of the three component-wise expressions $A_i^2 + A_i B_i + B_i^2$. Taking the square root after substituting the values from Table 2 with $a = b = 0.2$ gives $S'' = 2.59$. This number lies clearly between the one-teacher value $S = 4.19$ and the three-teacher value $S' = 1.12$, and it gives an early numerical illustration of the ordering that will be proved next.

We can now compare this variance with the variances for the other two arrangements. Using Eq. (15) and the definitions of P_A, P_B, and P_{AB} in Eq. (13), we obtain

$$S''^2 - S'^2 = \frac{2P_A + P_{AB} + 2P_B}{9}. \tag{23}$$

Similarly, using Eq. (8), we obtain

$$S^2 - S''^2 = \frac{2(2P_A + P_{AB} + 2P_B)}{9}. \tag{24}$$

As shown after Eq. (13), each of P_A, P_B, and P_{AB} is positive under the assumptions of the model. Both differences are therefore positive. Since standard deviations are non-negative, we obtain

$$\boxed{S' < S'' < S.} \tag{25}$$

Therefore, among these three arrangements, the variation for a given student is the smallest when three different teachers evaluate one component each, the largest when one teacher evaluates all three components, and intermediate when two teachers share the three components.

The three arrangements, their possible teacher assignments, and the corresponding numbers of cases are summarised in Table 5. The standard deviations shown in the

Table 5: Summary of the three evaluation arrangements

Arrangement	**Teacher assignment**	**Number of cases**	**Standard deviation**
One teacher evaluates all components	(T_i, T_i, T_i), where $i = 1, 2, 3$	3	$S = 4.19$
Two teachers share the components	One teacher evaluates two components; another evaluates the third, while the third teacher evaluates none	$3 \times 3 \times 2 = 18$	$S'' = 2.59$
Three teachers, one per component	All permutations of (T_1, T_2, T_3)	$3! = 6$	$S' = 1.12$

last column are obtained from Eq. (9) and by taking the square roots of the variance expressions in Eqs. (15) and (22), using the values of a and b given in Table 2, as used in Tables 3 and 4.

Table 6 lists the values of S, S'', and S' for different combinations of a and b, using the component marks and full marks given in Table 2.

Table 6: Standard deviations for the three evaluation arrangements for different combinations of a and b

a	b	S	S''	S'	S/S'
0.1	0.2	2.66	1.63	0.68	3.88
0.2	0.1	3.70	2.29	1.02	3.64
0.1	0.3	3.27	2.01	0.82	3.98
0.3	0.1	5.34	3.31	1.48	3.61
0.1	0.4	3.92	2.40	0.97	4.06
0.4	0.1	6.98	4.33	1.94	3.60
0.2	0.3	4.73	2.91	1.24	3.81
0.3	0.2	5.79	3.58	1.58	3.67
0.2	0.4	5.31	3.26	1.37	3.88
0.4	0.2	7.41	4.59	2.04	3.64
0.3	0.4	6.82	4.20	1.80	3.78
0.4	0.3	7.88	4.87	2.14	3.69

Note: The ratios S/S' are calculated using the unrounded values of S and S'.

3.2. Determination of a and b based on marks

Table 7 lists a hypothetical set of marks obtained by three students in an examination where C_1, C_2, and C_3 have been evaluated by T_1, T_2, and T_3, respectively. Table 7 contains only the marks actually recorded in the examination. The marks M_{12} and M_{13}, which T_1 would have awarded for components C_2 and C_3, respectively, are not available

from these observations and must therefore be estimated using an additional assumption.

Table 7: Marks obtained by three students in an examination (hypothetical)

	Marks given by T_1 for C_1 (M_{11})	**Marks given by T_2 for C_2** (M_{22})	**Marks given by T_3 for C_3** (M_{33})	**Total Marks**
Student 1	6.0	7.0	18.5	31.5
Student 2	7.0	8.5	20.0	35.5
Student 3	8.0	9.5	21.5	39.0

The marks that T_1 would be expected to give for C_2 and C_3, namely M_{12} and M_{13} respectively, are estimated from the marks given by T_1 for C_1. For the estimation of M_{12} and M_{13}, we additionally assume that the marks given by the reference teacher T_1 are proportional to the full marks of the three components of the paper, as follows:

$$\frac{M_{11}}{F_1} = \frac{M_{12}}{F_2} = \frac{M_{13}}{F_3}. \tag{26}$$

We use this proportionality assumption here only to illustrate the estimation procedure. In actual assessment data, the ratios M_{1i}/F_i need not be equal. Thus, for each student,

$$M_{12} = \frac{M_{11}}{F_1}F_2, \qquad M_{13} = \frac{M_{11}}{F_1}F_3. \tag{27}$$

For Student 1, for example, $M_{12} = (6.0/10) \times 15 = 9.0$ and $M_{13} = (6.0/10) \times 25 = 15.0$. Table 8 lists these values for all three students.

Table 8: Estimated M_{12} and M_{13} values for the three students

	M_{11} (from Table 7)	$M_{12} = \dfrac{M_{11}}{F_1}F_2$	$M_{13} = \dfrac{M_{11}}{F_1}F_3$
Student 1	6.0	9.0	15.0
Student 2	7.0	10.5	17.5
Student 3	8.0	12.0	20.0

From the expressions in Table 1, the marking parameters are obtained as

$$a = 1 - \frac{M_{22}}{M_{12}}, \qquad b = \frac{M_{33} - M_{13}}{F_3 - M_{13}}. \tag{28}$$

For the present data, $M_{13} < F_3$, so the denominator is not zero. For Student 1, these expressions give $a = 1 - 7.0/9.0 = 0.2222$ and $b = (18.5 - 15.0)/(25 - 15.0) = 0.35$. The same calculation is carried out for the other two students. Table 9 lists the values and their mean.

Table 9: Determination of a and b using Tables 7 and 8

	$a = 1 - \dfrac{M_{22}}{M_{12}}$	$b = \dfrac{M_{33} - M_{13}}{F_3 - M_{13}}$
Student 1	0.2222	0.3500
Student 2	0.1905	0.3333
Student 3	0.2083	0.3000
Mean	**0.2070**	**0.3278**

For the hypothetical data considered here, the mean values of a and b, shown in Table 9, provide illustrative estimates of the marking tendencies of teachers T_2 and T_3 respectively, taking T_1 as the reference teacher. The mean values are used here only as representative estimates for the illustrative dataset. They are averages across the three students, and they do not imply that a and b are the same for all students.

3.3. A Generalised Expression for Marks Given by a Teacher

Based on the model in Table 1, a generalised expression for the marks M_{xy}, given by a teacher T_x for the yth component, in terms of the reference marks M_{1y}, can be written as:

$$M_{xy} = (1 - L)(1 - a)M_{1y} + L\left[M_{1y} + b(F_y - M_{1y})\right]. \tag{29}$$

Here, L is either 0 or 1. Putting $L = 0$, corresponding to T_2, gives $M_{xy} = (1 - a)M_{1y}$. Putting $L = 1$, corresponding to T_3, gives $M_{xy} = M_{1y} + b(F_y - M_{1y})$. Thus, the term containing b is irrelevant for T_2, while the term containing a is irrelevant for T_3. Each non-reference teacher is therefore represented by only one relevant marking-tendency parameter: a for the lower-marking teacher T_2, and b for the higher-marking teacher T_3. Here, $L = 0$ represents a teacher who gives lower marks than T_1, while $L = 1$ represents a teacher who gives higher marks than T_1. The reference teacher T_1 is represented separately by the convention $a = b = 0$. Putting $a = b = 0$ in Eq. (29) gives $M_{xy} = M_{1y}$. This convention is used only to represent T_1, and it is not part of the positive-parameter assumptions for T_2 and T_3. The value of L is therefore not needed for T_1. The values of a, b, and L used for the three teachers are listed in Table 10.

Table 10: Values of a, b, and L for three teachers

	a	b	L
T_1	0	0	
T_2	0.2070		0
T_3		0.3278	1

Note: NA denotes not applicable. The values of a and b listed for T_2 and T_3, respectively, are the mean values obtained from the illustrative data in Tables 7–9.

The percentage deviation of the mark given by T_x from the reference mark M_{1y} is defined as

$$D_{xy} = \frac{M_{xy} - M_{1y}}{M_{1y}} \times 100. \tag{30}$$

Substituting Eq. (29) into Eq. (30), we obtain

$$D_{xy} = \frac{(1-L)(1-a)M_{1y} + L\left(M_{1y} + b(F_y - M_{1y})\right) - M_{1y}}{M_{1y}} \times 100. \tag{31}$$

Table 11 lists the values of D_{xy} for different combinations of the parameters a, b, and L.

Table 11: Values of D_{xy} for different combinations of a, b, and L

a	b	L	D_{xy} (%)
0.1		0	-10.00
0.2		0	-20.00
0.3		0	-30.00
0.4		0	-40.00
0.5		0	-50.00
	0.1	1	5.38
	0.2	1	10.77
	0.3	1	16.15
	0.4	1	21.54
	0.5	1	26.92

Note: NA denotes not applicable. For illustration in Table 11, $y = 1$ is considered with $M_{11} = 6.5$ and $F_1 = 10$. The value $M_{11} = 6.5$ is chosen arbitrarily for illustration and is not taken from the data in Table 2; the full mark $F_1 = 10$ is the same as in Table 2.

For $L = 0$, Eq. (31) reduces to $D_{xy} = -100a$. Thus, when $L = 0$, the percentage deviation does not depend on M_{1y}, as also shown in Table 11. For $L = 1$, the same equation reduces to $D_{xy} = 100b\left(\frac{F_y}{M_{1y}} - 1\right)$. In this case, the percentage deviation depends on both b and the ratio of the full marks to the reference marks for the component.

4. DISCUSSION

The analytical results in Sections 3.1.1 and 3.1.2 establish the ordering $S' < S'' < S$ under the conditions of the present model. Table 5 gives a compact summary of the three arrangements, the numbers of possible assignments, and the standard deviations for the illustrative example. Tables 3 and 4 illustrate the comparison between the first and third arrangements for $a = b = 0.2$: the standard deviation is $S = 4.19$ when one teacher evaluates all three components, whereas it is $S' = 1.12$ when three different teachers evaluate one component each. For the intermediate arrangement, taking the square root of Eq. (22) gives $S'' = 2.59$ for the same values of a and b. Thus, the numerical example follows the analytical ordering. Table 6 gives the values of all three standard deviations

for twelve combinations of a and b. In every case, the values satisfy $S' < S'' < S$. The ratio S/S' is also shown for comparison; its smallest value in Table 6 is 3.60. These results show that, for the same student and the three specified sets of assignments, the variation in the total is the smallest when three different teachers evaluate one component each, the largest when one teacher evaluates all three components, and intermediate when two teachers share the components.

When three different teachers evaluate the three components, the marking tendency of each teacher is confined to a single component. A reduction on one component can therefore be offset by an increase on another, so the higher and lower tendencies partly balance each other in the total. In the intermediate arrangement, the marking tendency of one teacher influences two components, while the tendency of the second teacher influences only one component. The imbalance is therefore larger than in the arrangement where three different teachers evaluate one component each, but it is still smaller than when a single teacher controls all three components. The resulting variation therefore lies between the variations of the other two arrangements, exactly as the analytical calculation shows.

The results suggest that, within the assumptions of the present model, using different teachers for different components can reduce the influence of any single teacher's marking tendency on the total. Thus, assigning different teachers to different components can lead to a less variable evaluation process, within this model.

The results in Tables 7–9 show that, within the assumptions of the present model, the marking tendencies of the teachers, represented by a and b, can be estimated from the marks obtained by students in an examination. First, the marks that T_1 would be expected to give for C_2 and C_3 are estimated from the marks given by T_1 for C_1. These estimated marks are then used, together with the actual marks given by T_2 and T_3, to calculate a and b. Thus, the marking tendencies of the teachers can be calculated in this way, using the marks obtained by students in an examination. In the present example, the mean values of a and b are 0.2070 and 0.3278 respectively.

The results in Table 11 give a separate illustration of how the marking tendency of a teacher can change the marks, relative to the reference teacher. For the lower-marking teacher, the percentage deviation is negative, and it increases in magnitude as a increases. For the higher-marking teacher, the percentage deviation is positive, and it increases as b increases. For the illustrative case used in Table 11, with $y = 1$, $M_{11} = 6.5$, and $F_1 = 10$, the deviation ranges from -10.00% to -50.00% for the lower-marking teacher, and from 5.38% to 26.92% for the higher-marking teacher. These results show how the parameters a and b represent the extent of deviation from the reference teacher in this model.

The comparison in this paper concerns the dispersion of total marks across the specified teacher-assignment cases. It does not, by itself, establish that the resulting marks are more accurate or valid, nor does it provide a psychometric reliability coefficient [4]. The result is conditional on the model assumptions, including the use of T_1 as the reference teacher and the equal weighting of the cases within each arrangement. An empirical

assessment of accuracy or reliability would require suitable student-performance data and an independent basis for evaluating the marks.

5. CONCLUSION

Within the assumptions of this model, the total marks of a given student vary the least when three different teachers evaluate one component each. They vary the most when one teacher evaluates all three components. The variation lies between these two cases when two teachers share the components. The reason is simple. When each teacher evaluates only one component, a lower mark on one component can be balanced by a higher mark on another. The study also shows how the marking tendencies can be estimated from the marks given by the teachers, within the model.

The model considers an assessment consisting of three separate components evaluated by three teachers. Although this arrangement is used here to develop the mathematical framework, the underlying approach is not restricted to continuous internal assessment. It provides a way for educators and administrators to examine how different evaluator assignments may affect the variation in total marks. The model can also be used as a teaching example in courses on assessment or educational statistics. The method for estimating marking tendencies shows how such tendencies can be studied using marks data, within the model.

However, this is a theoretical model with simplified assumptions. It takes one teacher as a reference for comparison. It assumes the marking tendencies are the same for all three components for a given student. It gives equal weight to all cases within each arrangement. It uses hypothetical numbers for illustration. For estimation, it uses a proportional marks assumption, which may not hold in real data. The result is about variation of totals across assignment cases only. It does not prove that shared evaluation is more accurate, more valid, or more reliable in a psychometric sense. It does not provide a generalizability coefficient or variance components. Real assessment involves many more factors. Readers should treat this as an illustrative framework, not as an empirical claim about actual assessment systems.

Future work can use real internal assessment marks to estimate marking tendencies. To do this without the proportional assumption, the same component would need to be evaluated by more than one teacher. Future work can also allow marking tendencies to differ across components or to vary across students. It can study how the magnitude of the variation changes with different distributions of full marks. It can extend the model to more than three components and more than three teachers. Future work should also examine assessment arrangements in which the number of evaluators is different from the number of components being assessed. Such cases may include situations where one evaluator assesses several components, where a component is assessed by more than one evaluator, or where the evaluators are distributed unevenly among the components. The model can be extended to study how different patterns of evaluator assignment affect the variation in total marks. It can test the framework with data from actual examinations conducted in any course of study. The framework may also be adapted for other multi-component evaluation settings.

REFERENCES

1. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. DOI: 10.1080/0969595980050102
2. Huysamen, G. K. (1990). The application of generalizability theory to the reliability of ratings. *South African Journal of Psychology, 20*(3), 200–205. DOI: 10.1177/008124639002000310
3. Brennan, R. L. (2001). *Generalizability theory*. Springer. DOI: 10.1007/978-1-4757-3456-0
4. Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine, 119*(2), 166.e7–166.e16. DOI: 10.1016/j.amjmed.2005.10.036
5. Atılgan, H. (2019). Reliability of essay ratings: A study on generalizability theory. *Eurasian Journal of Educational Research, 19*(80), 133–150.