An Empirical Classification Model for Errors in High School Mathematics

Article *in* Journal for Research in Mathematics Education · January 1987

DOI: 10.2307/749532

3 authors:

Nitsa Movshovitz-Hadar
Technion - Israel Institute of Technology

Orit Zaslavsky
Technion - Israel Institute of Technology

Shlomo Inbar

Journal for Research in Mathematics Education
1987, Vol. 18, No. 1, 3–14

AN EMPIRICAL CLASSIFICATION MODEL FOR ERRORS IN HIGH SCHOOL MATHEMATICS

NITSA MOVSHOVITZ-HADAR, *Technion—Israel Institute of Technology*
ORIT ZASLAVSKY, *Technion—Israel Institute of Technology*
SHLOMO INBAR, *Beth-Berl College, Israel*

A content-oriented analysis of written solutions to test items in Israeli high school graduation examinations in mathematics yielded a system of six error categories: misused data, misinterpreted language, logically invalid inference, distorted theorem or definition, unverifed solution, and technical error. This system enabled most of the documented errors to be classified. A reliability test showed that the categories are inclusive and mutually exclusive.

According to Radatz's (1980) historical survey, error analysis has been of interest to the mathematics education community for at least 70 years.

Research interest focused on five objectives:

1. Listing all potential error techniques;
2. Determining the frequency distribution of these error techniques across age groups;
3. Analyzing special difficulties, particularly encountered when doing written division, and when operating with zero;
4. Determining the persistence of individual error techniques;
5. Attempting to classify and group errors (p. 19).

Our study falls within the last category.

Much of the research on error analysis has been concerned with errors in arithmetic. In the last decade, however, students' errors in more advanced topics became a major source of information for the study of older students' difficulties in mathematics (Davis & Cooney, 1977; Hart, 1981; Quintero, 1983). A qualitative analysis of errors in high school mathematics was at the focus of our study.

GOALS OF THE STUDY

The repeated failure of a large percentage of high school students on the graduation examinations in mathematics in Israel motivated a systematic investigation of their common mistakes. We undertook this study to develop an empirical classification model for errors in high school mathematics and to demonstrate its reliability. *Empirical* means that the investigation relied solely on data in students' answer books for a comprehensive examination, assuming no theory to begin with. The only basic assumption was that most

---

This study was supported by the Israel National Academy of Science, Fundamental Research Foundation File 10-(VII)-4, and by the Israel Ministry of Education.

of the errors high school students commit in mathematics are not accidental and are derived by a quasi-logical process that somehow makes sense to the student. This assumption is an extension to high school mathematics of Ginsburg's (1977, p. 129) principles underlying errors in arithmetic.

Radatz (1979), having in mind the elements of information-processing theory, suggested a model that provides a classification of causes for errors by describing five mechanisms that yield errors across mathematical topics: errors due to language difficulties; errors due to difficulties in obtaining spatial information; errors due to a deficient mastery of prerequisite skills, facts, and concepts; errors due to incorrect associations or rigidity of thinking; and errors due to the application of irrelevant rules or strategies (pp. 165–168). We found this model inadequate for our study for reasons similar to those Radatz himself expresses:

> It is often quite difficult to make a sharp separation among the possible causes of a given error because there is such a close interaction among causes. The same problem can give rise to errors from different sources, and the same error can arise from different problem-solving processes. (pp. 170–171)

If this is true for his examples, which come from elementary mathematics, it is certainly true for errors in secondary mathematics.

Rather than turning to information processing, we sought a model anchored in an operations research analysis of the nature of errors. Later, we realized that the two approaches are not as distinct as we had thought. "To say that students process information is to use a doubtful metaphor, and how they process information is still the old question of how they learn" (Skinner, 1984, p. 949). Anyhow, the guiding idea in this study was to classify the errors by means of documented performance without appealing to processes in the students' minds that might or might not have yielded the errors committed and without faulting what the student did not do.

DEVELOPMENT OF THE CATEGORY SYSTEM

Source of Data

Data were collected from several samples (see below) of examinees' written answers to the Mathematics Matriculation Examination. This examination is prepared by a nominated committee and is administered nationwide at the end of the 11th grade (age 17) to all students in the nonscientific tracks of academic high schools in Israel on the completion of their mathematical education. About 20 000 students take this mathematics matriculation examination each year. All workbooks are transferred to the Ministry of Education for central grading that is conducted anonymously. Examiners are instructed to grade the solutions on a separate sheet, without leaving any comments on the workbook itself. The workbooks were made available to us after the grading process was finished.

Our investigation was repeated in two successive years. Each year the solutions to 18 open-ended test items were analyzed. The items covered the following topics: linear and quadratic functions, linear and quadratic equations, powers and logarithms, arithmetical and geometrical series, plane and solid geometry, elementary statistics, probability, and trigonometry. Sample items are presented in Table 1. Additional items appear as examples in the section describing the model.

Constructive Error Analysis

The errors were analyzed in a qualitative manner we called *constructive analysis*. It was conducted in the spirit of finding answers to the following questions: To what question (or questions) is the wrong solution a right answer? What logic can justify what the student in fact did? Only the student's performance as exhibited in his or her paper, including sometimes crossed-out, untidy parts, was considered in this process. We did not bother with questions like these: What did he or she not understand? Why did the student not do the right thing? Is the mistake a serious one? How far is it from the right solution?

Table 1
Examples of Items Analyzed in the Study

Topic	Item
Logarithms	a. $\log_4 x = 2.5$ Determine the value of x without tables. b. $\log x = 21 \log 3 + ½ \log 16 - \log 27$ Determine the value of x without tables.
Quadratic functions	The graph of $y = x^2 - 5x + 4$ is given in the figure. Line segment AB is perpendicular to the x-axis. a. What are the coordinates of E, C? b. $\overline{AB} = 1.75$ cm, calculate the measure of $\overline{OB}$. c. For what values of x will the function have negative values?

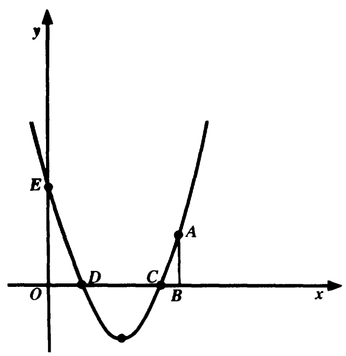
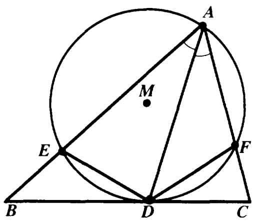
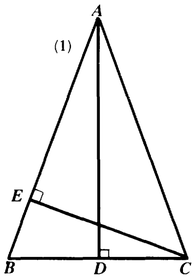
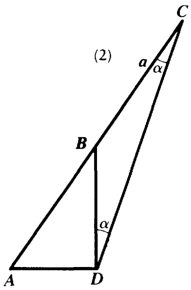


Table 1—*continued*

Topic	Item
Geometry	1. $\overline{AD}$ in the figure bisects $\measuredangle BAC$; circle M meets $\overline{BC}$ in point D. Prove that $BDE \sim DAF$.  2. ABC is a right-angle triangle ($\overline{AB} \perp \overline{BC}$). $\overline{AC} = 20$ cm, $\overline{BC} = 15$ cm. The plane ABC forms an angle of 30 degrees with another plane P, passing through $\overline{AB}$. What is the distance of C from plane P?
Trigonometry	1. In an isosceles triangle ABC ($\overline{AB} = \overline{AC}$), $\overline{BC} = 8.2$ cm, and the altitude is $\overline{AD} = 12.6$ cm (see figure). a. What is the measure of angle ABC? b. Find the measure of the altitude from C to AB.   2. In the figure the points A, B, C are on one line. $\overline{BD}$ is perpendicular to $\overline{AD}$. $\overline{BC} = a$, $\measuredangle BDC = \measuredangle BCD = \alpha$. State the measures of $\overline{DC}$ and of $\overline{AC}$ in terms of a and α. Calculate $\overline{DC}$ and $\overline{AC}$ if $\alpha = 16°10'$, $a = 10.5$ cm.
Probability	Twenty-five percent of the workers in a certain profession suffer from a disease associated with this work. Five workers are picked at random. a. What is the probability that at least 3 of the workers suffer from the disease? b. What is the probability that at most 2 of the 5 workers suffer from the disease?
Series	The manager of an industrial plant announced an award system. The total sum was \$19 250. The lowest award was \$500, and each one was larger than the previous one by \$250. a. How many awards were given? b. What was the highest award?

Formation of the Primary Error Classification System

One hundred ten examinees' answer books, selected at random from the pool of answer books in one year, were checked for incorrect solutions. For simplicity, only the first error was marked in each incorrect solution. The marking included "correct" solutions that arose from errors that canceled one another. However, errors derived from previous errors were disregarded. Errors occurring repeatedly—that is, in five or more answer books—were analyzed in the manner described above.

During the disputatious process of analyzing the errors, we realized that they were gradually starting to form clusters. We identified five error types: (a) the student added or ignored some specifiable datum, (b) the student translated a verbal expression into a mathematical expression that had a different meaning, (c) the student's inference was logically invalid, (d) the student applied an improper version of a definition or theorem, and (e) the student made a mistake in basic skills. This primary system included a description and examples for each category.

Reliability Check of the Primary System

To test the reliability of the primary system, we selected the 150 errors that occurred most frequently in the first sample. One occurrence of each error was clearly flagged in 40 of the 110 sample workbooks. Four experienced high school teachers of mathematics were asked to classify the errors according to written definitions of the five categories formulated earlier. A sixth category—"uncategorizable"—was added to make the system exhaustive. The coders were instructed to do their coding with respect to what the examinee wrote and not to any judgment of what the examinee missed, as teachers often do while correcting and grading students' work.

Of the 150 errors, 112 were identically categorized by at least three coders; 8 were classified as uncategorizable by at least one coder. To test the level of agreement among the four coders, Kendall's coefficient of concordance was applied to the rank orders of the category frequencies and found to be .91. Thus the hypothesis that the four ranks were independent was rejected at the .01 level of significance.

Revision of the Primary System

We concluded that the category system was inclusive for the set of items under investigation and did not find it necessary to add a category to the five main ones. We sought improvement, however, in the mutual exclusiveness of the categories themselves.

A modification of the category system took place following two inquiries:

1. An examination of the four coders' error classifications to find points of disagreement

2. A conference with each coder to discuss errors that were not classified the same way by all four coders, or were classified as uncategorizable

The results of the two inquiries did not indicate a need to split any category or to introduce any refinements in the category system itself. They did, however, suggest a re-examination of the criteria for classifying errors into categories. As a result, the definitions were revised to make them clearer, more precise, and more operational. For instance, we stated that an error should be classified as *distorted theorem or definition* only if the correct theorem or definition was identifiable. For *misused data* we stated that errors were to be included only if there was an explicit statement showing that the data were used. We noted that such errors could occur at any stage of the solution and not just at the beginning.

The modified category system was given to three additional experienced high school teachers of mathematics to apply to the same 150 errors. At this point, 88% of the errors were identically classified by us and by all three teachers. Only one error was marked as uncategorizable and that by only one teacher. Kendall's coefficient of concordance at this point reached the value .96. The errors on which there was still disagreement attracted our attention, and we later found that most of them arose from faulty editing (Movshovitz-Hadar, Zaslavsky, & Inbar, in press).

Using the modified category system, we drew a second random sample of 300 workbooks and established a classified inventory of 150 errors.

Application of the Revised System to a New Examination

While checking a sample of 110 workbooks from the examination for the following year, we realized that a new category was needed. In the first year, no item provoked an incorrect solution that in fact had no mathematical mistake but was "the right solution to the wrong problem." In the second year, one item provoked such an error. We repeatedly found solutions indicating that the student failed to realize that the answer did not satisfy the requirements posed in the question. Because we were unable to classify these incorrect solutions, we defined a new category and called it *unverified solution*.

At this point, we polished the statements of the category definitions and created a classified inventory of 130 second-year errors by applying the revised system to a random sample of 300 workbooks.

THE MODEL

The following six descriptive categories of errors are presented as a model for classifying errors in high school mathematics:

1. Misused data
2. Misinterpreted language
3. Logically invalid inference

4. Distorted theorem or definition
5. Unverified solution
6. Technical error

We present below each of the six categories described by its characteristic elements and followed by an example. Note that the characteristic elements are not necessarily independent. They are described in operational terms by telling what the student erroneously did as opposed to what he or she was supposed to do but did not do.

Misused Data

This category includes those errors that can be related to some discrepancy between the data as given in the item and how the examinee treated them. An explicit statement indicating that the given data were used must be included in the student's solution. Such an error may be committed either initially while putting the data together or later while processing the data. The main characteristic elements are as follows:

A: Designating as "given" a piece of information that neither is stated nor follows immediately from the given information. The student has added extraneous data.

B: Neglecting some given data necessary for the solution and compensating for the lack of information by explicitly adding irrelevant data.

C: Stating explicitly as a requirement (e.g., under "to be proved," "to be found," or "to be calculated") something that was not required in the problem.

D: Assigning to a given piece of information a meaning inconsistent with the text (e.g., using the height of a triangle in a solution to a problem that deals with the median).

E: Imposing a requirement that disagrees with the given information (e.g., forcing the properties of an angle bisector on an arbitrary line through the vertex of an angle).

F: Using a numerical value of one variable for another variable (e.g., using a given numerical value of distance as the numerical value of velocity).

G: Incorrectly copying some details from the test to the workbook.

Example: Given the series 1, 5, 7, which number should be added to each of the three elements to turn the series into a geometric one?

Incorrect solution:

"$7 = 1 + (3 - 1)d$

$6 = 2d$

$d = 3$"

Error analysis: (Letters in parentheses refer to the characteristics above.) The student imposed on the given series a property of an arithmetical series without verifying the middle term (A). The student also neglected the given information about the property of the target series (B). In addition, the student found a number, d, that was not the one required in the problem (C).

Misinterpreted Language

This category includes those mathematical errors that deal with an incorrect translation of mathematical facts described in one (possibly symbolic) language to another (possibly symbolic). The characteristic elements are as follows:

A: Translating an expression from natural language into a mathematical term or equation that represents a relation different from the one described verbally.

B: Designating a mathematical concept by a symbol traditionally designating another concept and operating with the symbol in its conventional use. (E.g., the problem deals with the sum of the last n elements in a series. The student designates this sum by S_n and later operates with S_n according to the formula for the sum of the first n elements.)

C: Incorrectly interpreting graphical symbols as mathematical terms or vice versa (e.g., wrongly matching an ordered pair with a point of intersection of two lines on a graph).

Example: A barrel full of wine weighs a kg; the empty barrel weighs b kg.
(i) What is the weight of m full barrels?
(ii) What is the weight of $2m$ half-full barrels?

Incorrect solution: A typical error in Part ii was "$2m\ \frac{a}{2}$."
instead of the right answer: "$2m\left(\frac{a-b}{2}+b\right)$."

Error analysis: The students translated the expression "$2m$ half-full barrels," namely, $2m$ barrels each half-full, into a mathematical term that represented "$2m$ halves of a full barrel," that is, m full barrels (Characteristic A).

Logically Invalid Inference

In general, this category includes those errors that deal with fallacious reasoning and not with specific content; that is, new information invalidly drawn from a given piece of information or from a previously inferred one. The characteristic elements are as follows:

A: Concluding from a conditional statement (if p, then q) its converse either in its positive form (if q, then p) or in its contrapositive form (if not p, then not q).

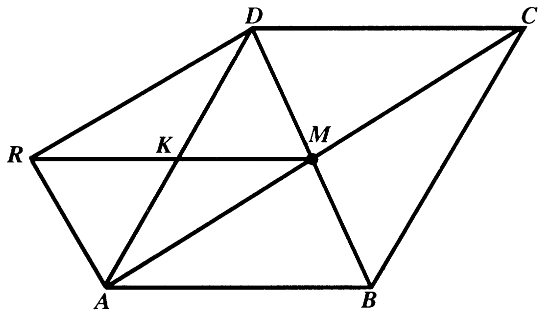
B: Concluding from a conditional statement (if p, then q) and from its consequent q that the antecedent p is valid; or concluding from a conditional statement and the negation of its antecedent (not p) that the negation of its consequent (not q) is valid.

C: Concluding that p implies q when q does not necessarily follow from p.

D: Using logical quantifiers such as "all," "there exists," or "at least" in the wrong place.

E: Making an unjustified jump in a logical inference; that is, stating that q follows from p without providing the necessary sequence of arguments leading from p to q, or providing erroneous arguments.

Example: $ABCD$ in the figure is a rhombus. Given that $\overline{AR} \| \overline{BD}$, $\overline{RD} \| \overline{AC}$, and K is the intersection of $\overline{RM}$ and $\overline{AD}$, prove that $AMDR$ is a rectangle.



Incorrect solution:

"$\measuredangle\ MAD \cong \measuredangle\ RDA$ (because $\overline{RD} \| \overline{AC}$);
$\measuredangle\ DAR \cong \measuredangle\ ADM$ (because $\overline{AR} \| \overline{BD}$); therefore $\measuredangle\ A \cong \measuredangle\ D$.
$AMDR$ is a rectangle because in a rectangle any two opposite angles are congruent and equal 90°."

Error analysis: The student made an unjustified jump from "$\measuredangle\ A \cong \measuredangle D$" to "$AMDR$ is a rectangle." The argument given, although consisting of a true statement, does not justify the jump (Characteristic E).

Distorted Theorem or Definition

This category includes those errors that deal with a distortion of a specific and identifiable principle, rule, theorem, or definition. The characteristic elements are as follows:

A: Applying a theorem outside its conditions (e.g., applying the law of sines, $\frac{a}{\sin \alpha} = \frac{b}{\sin \beta}$, where a and α do not belong to the same triangle as b and β).

B: Applying a distributive property to a nondistributive function or operation (e.g., $\sin(\alpha + \beta) = \sin \alpha + \sin \beta$; $\log \frac{a}{b} = \frac{\log a}{\log b}$; $(a + b)^n = a^n + b^n$.

C: An imprecise citation of a recognizable definition, theorem, or formula (e.g., in a parabola, the minimum is at $X_{min} = -\frac{b}{a}$ instead of $X_{min} = -\frac{b}{2a}$; $(a - b)^2 = a^2 + 2ab - b^2$).

Example: Same problem as for the previous category.

Incorrect solution:

"$\overline{DM} \| \overline{RA}$
$\overline{RD} \| \overline{MA}$

A quadrilateral with parallel opposite sides is a rectangle, therefore *AMDR* is a rectangle."

Error analysis: The student quoted a definition of a rectangle omitting the additional requirement on the angles (Characteristic *C*).

(For a deeper analysis of this category and for more examples, see Movshovitz-Hadar, Inbar, & Zaslavsky, 1986.)

Unverified Solution

The main characteristic of errors in this category is that each step taken by the examinee was correct in itself, but the final result as presented is not a solution to the stated problem. Had the examinee checked the "solution" against the requirements in the test, the mistake could have been avoided. It should be noted that very often students do not verify their results, but we cannot tell whether or not they did unless something in the solution is wrong.

Example: Given $i = \frac{nE}{nr + R}$, express n in terms of i, r, R, E.

Incorrect solution: The examinee ends up with the answer $n = \frac{i\,(nr + R)}{E}$, not noticing the n on the right-hand side. No error can be found in the process, but the presented solution is wrong.

Technical Errors

This category includes computational errors (e.g., $7 \times 8 = 54$), errors in extracting data from tables, errors in manipulating elementary algebraic symbols (e.g., writing $a - 4 \cdot b - 4$ instead of $(a - 4) \cdot (b - 4)$ but proceeding as if the parentheses were there as needed, which is a careless omission of parentheses), and other mistakes in executing algorithms usually mastered in elementary or junior high school mathematics (e.g., $\frac{71°}{2} = 35°5'$ instead of $35.5°$ or $35°30'$).

RESULTS

Altogether we dealt with about 280 errors—150 errors in the first year and 130 in the second. The distribution of errors among the categories of the model is shown in Table 2.

Table 2
Percentage of Errors in Each Category

Category	First year	Second year
Misused data	22	20
Misinterpreted language	17	18
Logically invalid inference	2	1
Distorted theorem or definition	34	32
Unverified solution	0	2
Technical error	25	27

One should not jump to conclusions with respect to the relative frequencies of errors. For instance, rather than concluding that students rarely commit logical errors, one might more reasonably speculate that the 36 items in the two examinations gave less opportunity to err this way than, say, to misinterpret the language. Balancing a test with respect to the opportunity to err in each category would be quite a challenge in itself; it would require the additional study of errors provoked by different items, perhaps in various styles. The model might contribute to an assignment of a level of difficulty to an item by the number of errors per category it provoked.

A classification of errors by their operational nature seems more promising than a classification by their causes because "a definite classification and hierarchy of error causes seems impossible to achieve" (Radatz, 1979, p. 171). For example, it is plausible to speculate that errors that we categorized as logically invalid inference (Category 3) were either due to incorrect association or rigidity of thinking (Radatz's Category 4) or were errors due to the application of irrelevant rules or strategies (Radatz's Category 5). On the other hand, errors due to a deficient mastery of prerequisite skills, facts, and concepts (Radatz's Category 3), which Radatz characterized as "deficits in . . . content and problem-specific knowledge . . . ignorance of algorithms . . . incorrect procedures in applying mathematical techniques" (pp. 165–166), might take, in actual performance, the form of either distorted theorem or definition (our Category 4) or of technical error (our Category 6).

The proposed model may help teachers foresee difficulties and obstacles and use this ability in planning their teaching so as to prevent as many of them as possible. Teachers may also use the model to identify a persistent tendency of individual students to make a certain type of error across several mathematical topics. This emphasis may suggest a specific remedial plan aimed at that type of error. Further research is needed to find out whether a certain teaching style or a specific cognitive style is associated with errors that fall into a particular pattern. Another possible application of the model is to provide a potential inventory of good distractors for multiple-choice items.

Moreover, descriptive models of mathematical errors, such as the model proposed in this article, if applied to a growing set of items and if solidly validated, can help one accumulate a classified inventory of errors, with indices of frequencies per category per item. Grouping items by the typical errors they yield and investigating common features of items in each group may prove indicative of errors expected from similar items, a process that may in turn give rise to a predictive model of errors in high school mathematics.

The category system presented in this article is an empirical one. It is the outcome of a content analysis of students' solutions to two samples of problems on a variety of mathematics topics. In considering the system, one should not overlook its limitations as a model. Even though it was validated for two specific sets of items, it undoubtedly needs much refinement before it can serve as a general model for errors in high school mathematics. On the other hand, it is solidly grounded in practice and therefore may prove useful for teachers, curriculum developers, and researchers interested in the diagnosis, remediation, and elimination of students' mathematical errors.

REFERENCES

Davis, E. J., & Cooney, T. J. (1977). Identifying errors in solving certain linear equations. *Two-Year College Journal, 7,* 170–178.

Ginsburg, H. (1977). *Children's arithmetic: The learning process.* New York: Van Nostrand.

Hart, K. M. (1981). *Children's understanding of mathematics: 11–16.* London: Murray.

Movshovitz-Hadar, N., Inbar, S., & Zaslavsky, O. (1986). Students' distortions of theorems. *Focus on Learning Problems in Mathematics, 8*(1), 49–57.

Movshovitz-Hadar, N., Zaslavsky, O., & Inbar, S. (in press). Sometimes students' errors are our fault. *Mathematics Teacher.*

Quintero, A. H. (1983). Conceptual understanding in solving two-step word problems with a ratio. *Journal for Research in Mathematics Education, 14,* 102–112.

Radatz, H. (1979). Error analysis in mathematics education. *Journal for Research in Mathematics Education, 10,* 163–172.

Radatz, H. (1980). Students' errors in the mathematical learning process: A survey. *For the Learning of Mathematics, 1*(1), 16–20.

Skinner, B. F. (1984). The shame of American education. *American Psychologist, 39,* 947–954.

AUTHORS

NITSA MOVSHOVITZ-HADAR, Senior Lecturer, Department of Education in Science and Technology, Technion—Israel Institute of Technology, Haifa 32000, Israel

ORIT ZASLAVSKY, Doctoral Candidate, Department of Education in Science and Technology, Technion—Israel Institute of Technology, Haifa 32000, Israel

SHLOMO INBAR, Mathematics Instructor, Beth-Berl College, Kfar-Saba 44905, Israel