WPI-CS-TR-2008-07 August 2008

An Analysis of Middle-School Math Errors Across Schools

by

Rob WEITZ[1], Neil HEFFERNAN[2], David ROSENTHAL[1] & Mingyu FENG[2]
[1] *School of Business, Seton Hall University, South Orange, NJ 07079*
[2] *Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609*
{weitzrob,rosentdv}@shu.edu
nth@wpi.edu

Computer Science Technical Report Series

WORCESTER POLYTECHNIC INSTITUTE

Computer Science Department
100 Institute Road, Worcester, Massachusetts 01609-2280

An Analysis of Middle-School Math Errors Across Schools

Rob WEITZ[1], Neil HEFFERNAN[2], David ROSENTHAL[1] & Mingyu FENG [2]
[1] *School of Business, Seton Hall University, South Orange, NJ 07079*
[2] *Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609*
{weitzrob,rosentdv}@shu.edu
nth@wpi.edu

Abstract. The little previous research comparing student errors across schools indicates that the distribution of students' systematic errors in one school does not significantly match those in other schools. We present here data from a set of middle-school mathematics problems, where student errors do seem to be comparable across the three schools studied. We conclude that student errors do seem to transfer for some problems, though not for others. One implication is that tutor remediation based on student performance at one school will likely be useful for students at another school.

Introduction

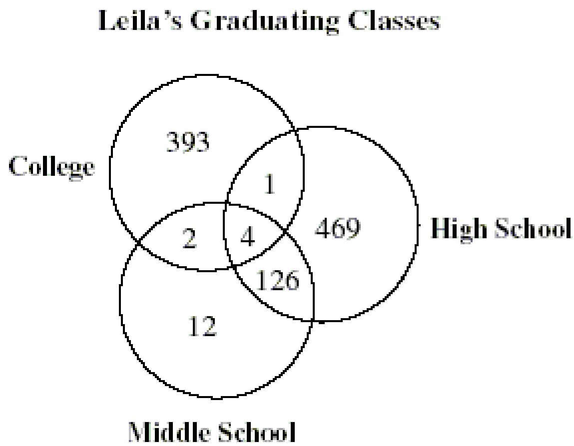
The little previous research comparing student errors across schools indicates that the distribution of students' systematic errors in one school does not significantly match those in other schools. This has been taken by some [1,2] to indicate that student errors do not ever transfer. The issue has practical implications if an intelligent tutor is to provide remediation based on systematic "bugs" in student problem solving [3].

In earlier research [4] we presented empirical data for one middle-school mathematics problem, where there appeared to be significant overlap of student errors across three different schools. In this paper, we present additional data, also from the domain of middle school mathematics. Our initial conclusions are that student errors transfer for some problems in this domain and not for others. We present sample problems in each category, and make some tentative hypotheses as to why this is the case.

Previous Research

The literature seems to indicate a single study devoted to the study of the transfer of student errors across institutions. Payne and Squibb [5] examined the errors made by 13-14 year-old students on algebra problems at three (English) secondary schools. One conclusion they drew (p. 455) is that, "the rules that do the most explanatory work in the three separate groups have surprisingly little overlap."

We previously [4] examined the responses of students at three schools to a single middle-school mathematics problem taken from the Assistment e-learning and e-assessing system [6]. The question is provided below in Figure 1.



The Venn diagram shows Leila's graduating classes from middle school, high school, and college. How many students graduated together from **both** Leila's middle school and high school?

Figure 1: The Middle-School Math Question Used In [4]

The correct answer is 130. However, the most common answer, provided by approximately 35% of students at each school was 126. Put another way, of the students who get this question wrong, approximately half got it wrong this way, at each school. Further we found that two of the schools shared the same top five most occurring responses, though in slightly different orders. These five errors accounted for some 2/3 of the incorrect responses at each of these schools. Four of these five errors appear among the top five for the third school, and in this case these four also comprised approximately 2/3 of the incorrect student responses. We get similar results using the top five errors of the third school to describe the results at the first two schools. These results seem to be in sharp contrast with those obtained by Payne and Squibb [5].

(On the other hand, the distribution of student errors was highly positively skewed – that is there were many rarely occurring errors. The skewness of the responses concurs with previous work in the area [5, 7]. We note as well that the commonality of responses across schools seemed to break down after the sixth response.)

In this paper we consider a set of additional problems, also middle-school mathematics problems taken from the Assistment e-learning and e-assessing system.

Data

We initially sampled 40 questions from the Assistment system system that reflected the categories of questions given in the Massachusetts Eighth Grade Math Assessment test. We ultimately had to remove 13 questions from this set as it included responses from students who had not yet covered the material related to the questions. The data we ultimately used are from eighth grade students, working problems from the seventh grade curriculum, as this assured us that the students had been exposed to the course material prior to answering the questions.

Of the 27 questions in our study, 14 showed demonstrable transfer of student errors. The others showed little to no transfer. (These numbers should be used as approximations as several questions were variants of the same basic problem.)

We first give detailed data on one of the questions for which student error transfer was evident. This will provide insight into the analysis process used in this project. We then provide summary descriptions and results for other questions. Finally we take a look at some problems where there was no apparent transfer.

Detailed Analysis of a Sample Question

A detailed analysis of a problem used in this study follows.

Consider the following problem:

25% of what number is 100?

The correct answer is 400. The distribution of student errors for this problem at the three schools we studied is presented below in Table 1.

Table 1: Top Ten Most Frequently Occurring Student Responses for a Sample Problem

	School ID				School ID				School ID		
Answers	1			Answers	2			Answers	3		
	# responses	%	cum %		# responses	%	cum %		# responses	%	cum %
400	153	49.35%		**400**	30	46.15%		**400**	37	46.25%	
25	48	30.57%	30.57%	25	19	54.29%	54.29%	25	22	51.16%	51.16%
4	47	29.94%	60.51%	4	6	17.14%	71.43%	4	6	13.95%	65.12%
75	23	14.65%	75.16%	75	6	17.14%	88.57%	75	3	6.98%	72.09%
2500	6	3.82%	78.98%	2500	2	5.71%	94.29%	0.75	2	4.65%	76.74%
100	4	2.55%	81.53%	0.25	1	2.86%	97.14%	2500	2	4.65%	81.40%
50	4	2.55%	84.08%	125	1	2.86%	100.00%	20	1	2.33%	83.72%
5	3	1.91%	85.99%			0.00%	100.00%	125	1	2.33%	86.05%
20	3	1.91%	87.90%			0.00%	100.00%	175	1	2.33%	88.37%
250	2	1.27%	89.17%			0.00%	100.00%	1	1	2.33%	90.70%
-999999	65			-999999	21			-999999	13		
Total Responses	375				86				93		
Total - Hints	310				65				80		
Total Incorrect	157				35				43		
Unique Responses	25				8				15		

- For each school, the ten most frequently occurring student responses are listed in the “Answers” column.
- The answer “-999999” corresponds with the student requesting a hint prior to attempting to answer the question. We exclude this data from consideration in our calculations for this study.

- The number of students who provided each answer is given in the "# responses" column.
- At the bottom of the table, the total number of responses to the question at each school is given (this is the same as the number of students who attempted to answer the question).
- In our calculations we only considered students who did not ask for a hint; this figure is provided in the row "Total – Hints."
- The row "Total Incorrect" gives the number of incorrect responses (the number of students who got the question wrong). It is determined by subtracting the number of correct responses from "Total - Hints." Here this figure is 310 – 153 = 157.
- The number of unique (different) responses is provided in the last column.
- The column labeled "%" provides two different calculations.
    - In the first row (where the answer is correct) it gives the percentage of all (non-hint) students who got the problem right (153/310 = 49.35%.
    - Where the answers are incorrect, it provides the percentage of students who provided that particular incorrect answer out of all incorrect responses. So, for (incorrect) answer 25, the percentage figure is 48/157 = 30.57%. Put another way, almost 1/3 of all the students who got this question wrong, got it wrong by answering "25."

We can see that the most commonly occurring wrong answer, 25, is the same at all three schools. Some 30% of students who answered the question incorrectly at school 1 answered it incorrectly this way. Likewise, almost 55% of students at school 2 and just over 50% of the students at school 3 fall into this category. The second and third most common wrong answers, 4 and 75 are the same at all schools. These three errors account for roughly 75% of all the wrong responses at school 1, 88% at school 2 and 72% at school three. It would seem that for this question students are making systematic errors across schools.

For example it appears that students who arrived at the answer 25 did so by calculating 25% of 100. To be clear, this can't be proven without collecting and examining the details of student solutions. However, it seems that in many cases we examined, targeted responses may be productively provided – in this case, for example with a remark to the student as to what he/she has done and why it's incorrect. If nothing else, such "knowledge" on the part of an intelligent tutor would inspire more respect on the part of the student than a generic hint or the worked out solution.

Brief Descriptions of Other Problems with Error Transfer

We analyzed all 27 problems as above. Here we briefly present some representative problems along with summary student data that demonstrates transfer of student errors.

400

The Eliot School plans to have the school library carpeted. The room is in the shape of a rectangle and measures 12 feet by 36 feet. If the carpet costs $32 per square yard including installation, how much will it cost to have the library carpeted?

The correct answer is $1536. The most common answer across all three schools, accounting for some 20% of the errors in each case, was $13,824. Clearly students neglected to note, or didn't understand, that the price is per square yard and the measurements are given in feet. Here again, targeted remediation is possible.
Other geometry oriented problems demonstrated similar characteristics that are arguably predictable. These include, in addition to neglecting different units, failing to include some part of the object in question, or failing to finally multiply by the price per square unit.

842
*Compute 8 – (-5 + 3 * 7)*

The correct answer is -8. The most common incorrect answers across all three schools were, though in the same order, 8, 22 and -6, comprising roughly between 35% and 45% of the errors. Other variations of this problem yielded similar results, where students demonstrate confusion in the precedence of mathematical operations, or simply got the sign wrong in their final answer. Targeted remediation is easy (and in our view, predictable) in these cases.

2264

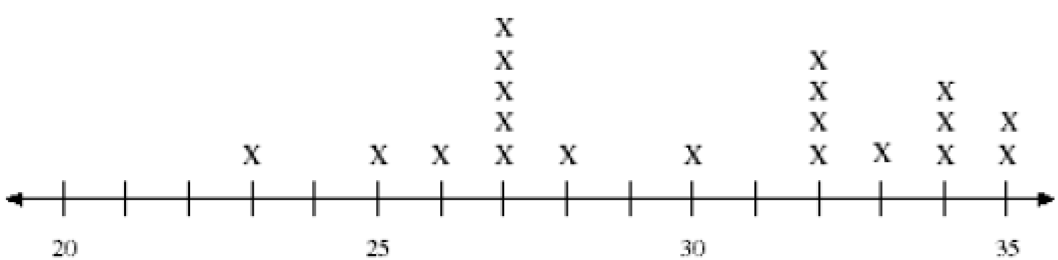
To sell their house, the Fords placed an advertisement in the local newspaper. For one week, they recorded the number of calls they received each day in response to the advertisement in the following chart.

S	M	T	W	Th	F	Sa
2	5	2	8	10	4	11

*What is the **mean** number of calls for the days shown in the chart?*

Here, eight is the most common wrong answer at one school, and the second most common wrong answer at the other two. It appears that students confuse median with mean. (Six is the correct answer.)

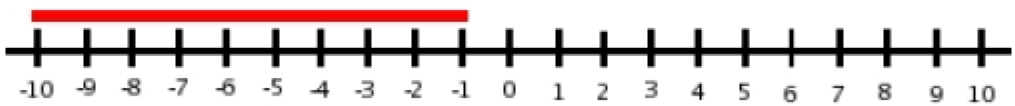
2267



The line plot above represents the number of raisins that Janika's class counted in each of 20 boxes of cereal. What was the median number of raisins in a box?

The correct answer is 31. The most common incorrect answer for the first two schools was 30. This was the second most common answer in school 3. It accounted for between 20% and 28% of the errors. Apparently students don't know that when there are an even number of observations, the median is obtained by averaging the middle two. (But they do know something about the median being a middle value.)

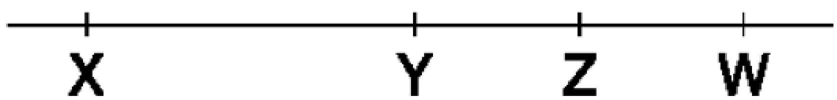
525



How many units apart are -10 and -1 on the number line?

The correct answer is 9. The top three incorrect answers at all three schools were 10, 8, and -9. These accounted for between 89% and 100% of student errors.

178



In the figure above:
distance *yz* equals distance *zw* and
distance *xy* equals distance *yw* and
distance *yz* equals 4

What is the value of distance *xw*?

The most common incorrect answer across all three schools is 8, accounting for about 50% of incorrect answers. The correct answer is 16.

Problems With Little to No Transfer of Student Errors

About half the problems we examined did exhibit common student errors and/or common student errors across schools. For some we hypothesize the problems were simply too difficult for the students. We continue study this phenomenon. However thus far we have identified certain basic computation problems as typically not exhibiting regular student errors. The MCAS includes basic multiplication of numbers with fractional components. We examined the following two problems:

2711
What is ¾ of 1 ½?

1163
*Compute 43.68 * 2.5 = ?*

In each case most students got the problems correct. In the second problem there were a number of incorrect answers due to rounding errors or errors in placing the decimal point. In any case, the incorrect responses seemed to vary greatly within and between schools. In our view, this coincides with the results of Payne and Squibb [3] who focused on algebra problems, and with similar work documenting student errors on subtraction problems [7].

Discussion

It appears that for some problems there are systematic student errors and that these errors tend to transfer between schools. This is an intuitive result for anyone who has ever been part of a team grading student exams: typically there is a standard rubric for taking away points on individual problems for errors that students are expected to make. However, researchers continue to point to Payne and Squibb, who in our view, showed that student errors do not transfer for the domain they studied. Generalizing to all domains is inappropriate.

Our feeling is that in some cases, the common errors we've come across here are due to a real misunderstanding of how to solve the problem, some are due to errors of attention (or "slips") and some are due to simple guessing strategies. In some cases, it may be difficult to differentiate these. However, we believe there is real value in pointing out to a student, for example, that he/she has arrived at an incorrect answer for a mathematics problem as a result of failing to follow the rules of precedence, or has gotten a geometry oriented problem incorrect due to forgetting to include two of the sides, or in another problem has forgotten to convert to common units.

References

[1] Mitrovic, A., Mayo, M., Suraweera, P. and Martin, B. (2001), Constraint-Based Tutors: A Success Story. In L. Monostori, J. Vancza and M. Ali (Eds.), *Proceedings of the 14th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE – 2001)*, 931-940, Berlin Heidelberg New York: Spring-Verlag.

[2] Kodaganallur, V., Weitz, R. R. and Rosenthal, D. (2006). An Assessment of Constraint-Based Tutors: A Response to Mitrovic and Ohlsson's Critique of "A Comparison of Model-Tracing and Constraint-Based Intelligent Tutoring Paradigms" *International Journal of Artificial Intelligence in Education*. 16, 291-321.

[3] Koedinger, K. and Anderson, J. (1997). Intelligent Tutoring Goes to School in the Big City. *International Journal of Artificial Intelligence in Education*, 8, 30-43.

[4] Weitz, R., Heffernan, N, Kodaganallur, V. and Rosenthal, D, "The Distribution of Student Errors Across Schools: An Initial Study," *Proceedings of the 13th International Conference on Artificial Intelligence in Education*, Los Angeles, CA, July 9 – 13, 2007,

[5] Payne, S.J. & Squibb, H.R. (1990). Algebra Mal-Rules and Cognitive Accounts of Error. Cognitive Science, 14, 445-481.

[6] Feng, M., Heffernan, N.T., & Koedinger, K.R. (2006). Predicting State Test Scores Better With Intelligent Tutoring Systems: Developing Metrics To Measure Assistance Required. In Ikeda, Ashley & Chan (Eds.) *Proceedings of the 8th International. Conference. on Intelligent Tutoring Systems*. Springer-Verlag 31-40.

[7] Van Lehn, K. (1982) Bugs Are Not Enough: Empirical Studies of Bugs, Impasses and Repairs in procedural Skills, *Journal of Mathematical Behavior,* 3, 2: 3-71.