The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
Paper ID #42112

Analyzing Grading Criteria for Linear Graphs: Implications for Advanced Mathematical Learning

Dr. Xiaojin Ye, State University of New York, Farmingdale

I am currently an Assistant Professor in the Department of Computer Systems at Farmingdale State College,SUNY. I earned my Ph.D. in Computer Science from the Graduate Center, CUNY. Additionally, I hold two master's degrees in mathematics, one from the Graduate Center, CUNY, and another from Kansas State University. My primary research interests include theoretical computer science, set operators, mathematical morphology, and mathematical education.

Prof. Carlos William Castillo-Garsow, Eastern Washington University

©American Society for Engineering Education, 2024

Analyzing Grading Criteria for Linear Graphs: Implications for Advanced Mathematical Learning

Xiaojin Ye
Department of Computer Systems
Farmingdale State College, SUNY

Carlos William Castillo-Garsow
Department of Mathematics
Eastern Washington University

Abstract

We conducted research to identify what features of a graph are important for college teachers with the intention of eventually developing a system by which a machine can recognize those features. Eleven experienced college algebra graders of a large state university were asked to grade graphs of linear equations generated by students in their classes, and interviewed to clarify what features of the graphs were important to them in grading. When grading each graph on a scale of 10 points, the graders generally agreed on the relative worth of particular features: a correct slope was worth 4 points, y-intercept was worth 4 points, labeling is worth 1 point. After that, and everything else was a matter of 1 point. Furthermore, the graders judged slope and intercept from two points (the y-intercept and the first point to the right). Returning to the students' work, the researchers saw that the students also placed extra importance on points to the right of the y-axis.

This grading style may reinforce students' thinking about only about two points in a line. Students understand graphing a line as just plotting two points and then connecting them. Beginning at the y intercept, students then go "over one and up m" to graph a line with slope m, or decomposing a fraction into "rise over run." In both cases, a slope is thought of as being composed of two discrete points. This method is good enough to generate graphs of lines from equations, but begins to fail as students begin generate equations from points. As mathematics becomes more complex, a strong foundation of continuous reasoning becomes even more necessary. We conclude that use of this grading style may have implications for student learning of more advanced mathematics. If a machine is doing the grading style, it can look at just those two points without making more work for the teacher. However, based on our research, it shows that replicating human grading may not be the best use of machine grading.

Keywords

grading, graph, features of a graph, linear function, algebra, grading criteria, student learning, machine grading

How College Algebra Teachers Grade Graphs of Lines

Distinguishing between discrete and continuous systems is critical for both mathematics and learning mathematics [5]. In the study of dynamical systems, chaotic behavior occurs in simple one-dimensional discrete systems [10]. However, the Poincaré-Bendixon theorem proves that chaos is impossible in continuous dynamical systems of fewer than three dimensions. In learning, it has been demonstrated that approaching a system from a discrete perspective can interfere with a student's learning of elementary differential equations [2,4].

Comprehensive literature reviews of graphing activities show that secondary students tend to focus on individual points and have difficulty interpreting global features of graphs, or features over intervals [6,9]. For example, Kerslake [8] asked 600 students ages 12-15 the question "Are there any points on the line between the points (2,5) and (3,7)? and if so, how many?" At every grade level, the majority of students identified a small finite number of points such as "5" or "10." Bell and Janvier [6] noticed that when students were asked questions that required interval answers such as "when is population B greater than population A," students tended to give point answers such as the maximum of B, rather than intervals. Castillo-Garsow [3] showed that students have difficulty distinguishing between situations that are continuous and linear from discrete situations (such as making regular payments) that are better modeled by a step function, because they are not attending to values in the intervals between marked points. Leinhardt et. al. [9, p. 11] concludes in part that "Overemphasizing pointwise interpretations may result in a conception of a graph as a collection of isolated points rather than as an object or a conceptual entity."

These difficulties also extend to high school teachers. Thompson [13] identified the case of "Sandra," where a high school teacher struggled to find the y-intercept of a line from two given points (3,1) and (7,4). In this case, Sandra thought of the slope in terms of "over 4 and up 3." However when changing by -4 from (3,1), she passed the y-axis, and didn't know how to adjust for a change in x of -3, rather than -4.

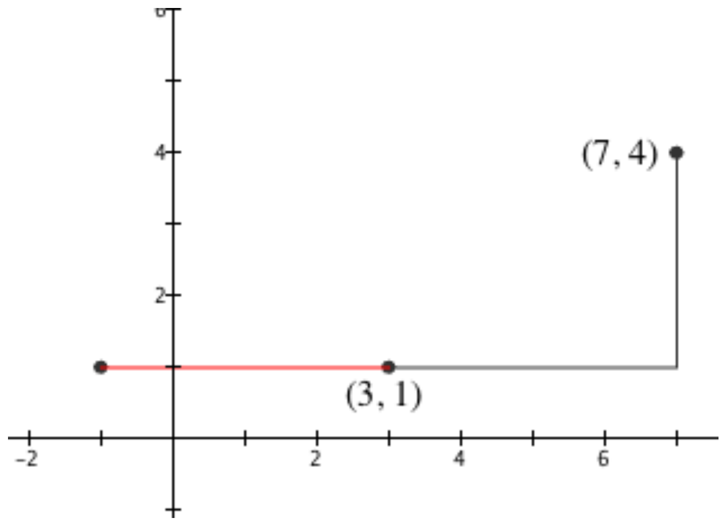

Figure 1. Sandra's board work for finding the y-intercept of a line from two points (reproduced from [13]). By going "over 4," Sandra was unable to land on the y-axis, and could not find the y-intercept.

This research project began initially from interest in streamlining the grading process for large classes by developing a method by which a computer could automate grading of hand drawn graphs. The difficulty in the design was that hand drawn graphs are symbolic and communicative rather than precise. Hand drawn graphs are sketches based on the cultural expectations of students and teacher. We were concerned that the more obvious quantitative methods such as linear regression might not capture the communicative aspects of graphing. With an eye to this problem, we designed a series of interviews to identify the features of a graph are important for college teachers so that in the future a machine can recognize or improve on those features. We elected to study college algebra teachers, because college algebra is a basic math course and enrolls over 1000 students each year at the university where the study took place. Therefore, developing machine grading for college algebra had the potential to benefit a large population. The details of the interviews and results can be found in [1].

However, in the process of discussing the results of these interviews, we found that our focus shifted. The standard that the graders in our study used seemed to reinforce well-documented difficulties that students have in graphing described above [6,9]. The purpose of this article is to summarize the results of this research and explore the consequences of those results. It may be that machine grading, by way of saving time, also creates new possibilities for grading standards that may better serve the mathematical needs of students.

Method

The grader pool for the interviews was composed of 11 graders: one undergraduate grader, nine graduate teaching assistants in mathematics, and one mathematics professor, who were selected to be representative of the whole grading system for the college algebra course at a large public university. The differences level of education did not appear to affect the results. Although there were differences in grading styles within the graduate students, the professor and the undergraduate grader were very similar to the majority. All these chosen college algebra graders were native English speakers and had been educated in America. For privacy, the graders were all given pseudonyms.

The interview was a clinical interview [7] with two parts. The first part was a task-based interview with a think-aloud protocol, in which the graders were asked to assign grades to student graphs. The second part was semi-structured discussion, in which the graders reasons for assigning graders to particular graphs were explored. Each one-on-one interview took 30 to 45 minutes and was voice recorded.

Each grader was asked to grade twenty graphs by using the grading scale of 0 to 10. The graphs were genuine, anonymized student solutions to the tasks of graphing two equations: nine graphs of the equation y+3=x-2 (indexed alphabetically as Graphs A-I) and eleven graphs of the equation -4y=2x+1 (indexed numerically as Graphs 1-11). The student solutions were taken from earlier one-on-one student interviews of college algebra students in the same semester.
These student solutions were presented in bulk, so that the graders could grade more efficiently during the interview, as well as enter the rhythm that they had when grading a real problem set. For realism, all the chosen graders were given the equations a week before their interviews so that they could have some time to think about how to grade the problems. All the graphs were shown to the graders in the same order.

Prior to the interviews, the interviewer scored all the graphs herself, with the intention of discussing with the graders and graphs where their score differed substantially from hers. After the graders finished all their grading, the interviewer had a short discussion with them about any graph where the grader's score differed substantially from the interviewer's, followed by discussions of up to seven pre-selected graphs (depending on time available). During these discussions, the interviewer pointed out features of these graphs that she suspected the graders had not noticed, and gave the graders an opportunity to change their grade, in order to see how important each feature was to the grader. Both the interview and the retrospective analysis took the general form of a grounded model construction study [7].

Results

There was a high level of agreement among the graders. For problems scored out of 10 points, the standard deviations for most problems were lower than 2. There were only three standard deviations bigger than 2. This larger standard deviation was the result of a single grader. In all these 3 graphs, Daniel gave much lower grade than other graders. In terms of effect on scoring, the grading criteria of all these graders were similar to each other.

In the retrospective analysis of the interviews, student responses that we call F, G, 5 and 9 proved to be particularly useful in highlighting individual grading choices. Each graph had unique errors (such as correct slope and incorrect y-intercept) that were particularly useful in identifying the relative importance of features to graders. These graphs, and the discussions that graders had with the interviewer about them are the primary result of this paper.

Slope

Student solution F was a very useful graph from which to obtain the graders' ideas of how important the slope is in graphing (Figure 2). In this graph, the student made an error in calculating the y-intercept, so all the points on the graph are incorrect, but the slope between the points is approximately correct (although somewhat sloppily drawn).

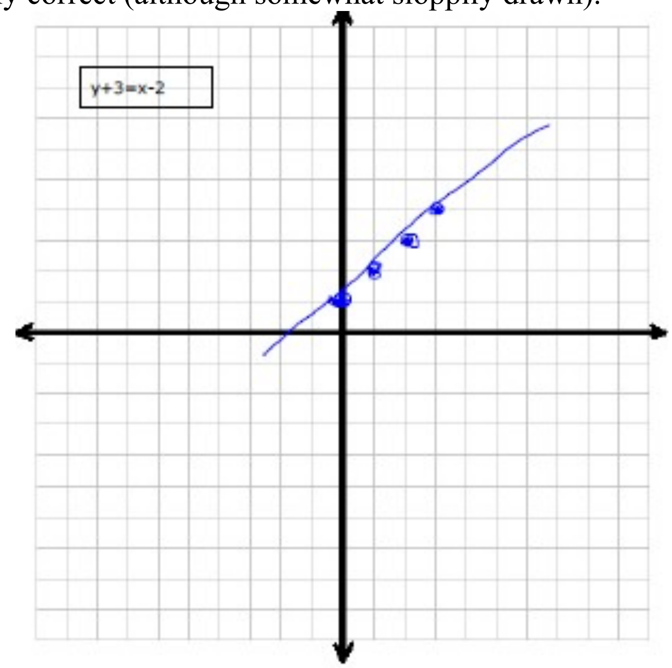

Figure 2. Student solution F, for y+3=x-2

Before discussion with the graders, the mean was 4.09, the highest score was 7 and the lowest score is 1. The reasons for giving these grades are below:

In the interview, Daniel gave the lowest grade, 1 point for this graph. He said, "The graph is completely wrong and no information telling me what they did. When I was in class I told my students what I expected of them; if you only gave me a graph completely wrong I only gave them 1 point." After the interviewer told him the slope was very close to correct, he said, "If the student labels the x and y intercepts so that I would see what exactly is the slope I may gave them 3 to 4 points."

Ray gave 3 points for this graph, and judged the slope of the line to be correct. However, he would have liked the student to show more work in order to get the full points from a correct slope. Gerry also gave this graph 3 points, he said, "even if is a right slope but he did not write down what the equation was y equals x minus 5; if they write that down, we know the slope should be 1." We see that for Ray and Gerry a correct slope alone is worth 3 points, with potentially more for showing work of some kind to establish that the student drew a slope of 1 intentionally.

Frank and Alice had clear grading standards for slope and gave 5 points for this graph.

Sandra said, “A correct slope is awarded 4 points.” Jimmy said, “The student can have 2 points for having the correct slope.” Dolly said, “I will give about half of the points for the slope, and I would give 5 or 6 for this slope depending on the student’s actual work.” She gave 6 points to this graph.

Kara awarded 4 points to this graph originally, but after the interviewer mentioned that only the slope was correct in this graph, she said, “I usually gave 2 points for a correct slope only, so I would like to change the grade to 2 points.”

Michael gave 5 points for a correct slope, and after discussion he realized that the graph did not label axes, he changed the grade to 4 points.

Normally Sam gave a correct slope 5 points, but he gave 7 points for this graph, because he judged that the student had made an algebra mistake in solving for y, adding 3 to 2 to arrive at the equation y=x+1. He stated that thought the student understood how to graph and he could not take off a lot points. He was the only grader to mention this potential algebra mistake.

After discussion, the mean was 3.74, the highest score was 6 and the lowest score was 1. However, from the interviews, we see that the value given to a correct slope was higher than the average grade given. First, Daniel gave 1 point for this graph which was very low compared to other graders, and he said would give 3 or 4 points for a correct slope with more clear descriptions to show the exactness of the slope. Both Ray and Gerry said that the value of a correct slope should be higher than the 3 points that they gave for this graph. Michael and Sam gave a clear signal that they would give 5 points for a correct slope, and they changed the grade after discussion because the other errors of this graph, but not the slope, such as lack of labeling.

If as they said, Ray and Gerry gave 4 points for a fully correct slope, and Daniel gave 4 points for a correct slope, we had 5 graders who gave 5 points, 4 graders who gave 4 points, and 2 graders who gave 2 points for a correct slope. That means the mean of the correct slope is 4.09. Therefore, roughly speaking, a correct slope was worth 4 points, with very close agreement.

Intercepts: y-intercept; both x- and y-intercepts

Two common ways of graphing a line form an equation are graphing the y-intercept and a point to the right, or plotting the x-and y-intercepts. For this reason, the intercepts, and the y-intercept in particular, are culturally important aspects of the graphing of a line.

A particularly useful graph for studying the value of intercepts was student solution G (Figure 3). The student created this graph by means of plotting the y-intercept and the point (1, -4), and then connecting the dots with a line. However, the student drew the line with a curve that resulted in the x-intercept being 4 instead of the correct x-intercept of 5.

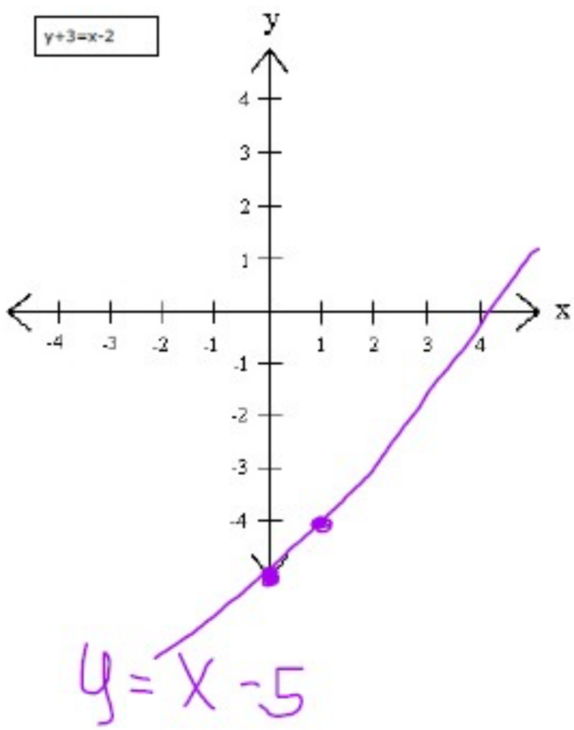

Figure 3. Student solution G, for y+3=x-2

Daniel gave 3 points for this graph and the other graders gave 7 to 10 points. In order to get more precise grade from the whole grading pool, we omitted Daniel's grade from the analysis of this graph. The mean score of this graph was 7.82, with the outlier omitted, the mean grade was 8.30. The graders gave the following reasons:

Gerry and Frank gave 10 points for this graph because they thought getting the right y-intercepts and a second point (or slope) is good enough to graph the equation and claimed it is more important that we can see this student knows how to graph it. After discussion, Gerry took off 1 point for curving the straight line.

Sandra said, "I gave 4 points for a correct slope, 4 points for the correct y-intercept, and 2 points for the whole correct graph, and I gave 1 point for the correct graph here because the curving of the line." So Sandra gave 9 points for this graph.

Kara said, "Two points can make up a line but it curved up and got the totally wrong x-intercept, and the x-intercept is important." She graded based on both the x and y intercepts, and took off 2 points for the incorrect x-intercept.

Sam said, "Some people said getting x and y intercepts, some people said getting the y-intercept and a point. Finding the y-intercept and a point are basically fine for me." He took off 2 points because of the line was not being drawn perfectly, not because of the x-intercept being wrong.

Alice gave 5 points for a correct y-intercept. She said, "In this graph the y- intercept is correct, but the student makes a mistake of the slope by having the x-intercept go to 4 instead of 5. I will give the student 8 points because it looked the student start with a correct slope by a point but somehow curving the line." She took off 2 points for the graph.

Michael said "The student got the y-intercept correct but he drew the line poorly and the x-intercept was wrong, I gave 7 points." When the interviewer suggested that 2 points could determine a line; he said he agreed but that he also paid attention to the x-intercept.

Ray and Jimmy attended to both x- and y-intercepts when grading all of the student solutions. They both took off 3 points for the incorrect x-intercept.

Dolly said, "I gave 5 points for the correct y-intercept, and the x-intercept here could be placed more accurately." So she gave 9 points for this graph. After discussion, she decided the line was curved and she changed her grade to 7 points.

The mean score was 7.55 after discussion. If we omit the grade of Daniel, the mean score was 8.00. On average, graders only took off 1 to 2 points for an incorrect x-intercept, and/or 1 to 2 points for the curve of the sketch. It is also possible that taking off points for the x-intercept could be considered taking off some points for the slope, because if both the *x* and *y* intercepts are correct, the slope of the line must be correct, however in the case of graph G, this is unlikely, as the marked point at (1, -4) established the slope. All the graders valued the y-intercept highly, and 8 out of 11 of them had a clear grading standard of giving 2 to 5 points for getting a correct y-intercept. Of the graders that had a clear standard, Gerry, Frank, Alice, Dolly said that they would give 5 points for getting the correct y-intercept, Sandra gave 4 points, Ray and Jimmy gave 3 points, and Kara gave 2 points. Therefore, a correct y-intercept was generally awarded about 4 points.

Prioritizing the Right Side of the Graph

Every straight line can be represented by an equation: $y = mx + b$, where m is the slope, b is the y-intercept. As discussed above, the graders gave greatest importance to the slope and the y-intercept in their grading. In grading these graphs, many graders used the y-intercept and a close-by point to determine whether the slope was correct. We found that in judging slope, the graders preferred to look at a close-by point on the right side of the y-intercept. In other words, the graders paid more attention to the right side of the graph than the left.

In Graph 5 (Figure 4), the student plotted exactly three marked points: the y-intercept (0, -1/4), a point to the left (-1/2, 1/4) and a point to the right (1/2, -1/2). The y-intercept and the point on the right were correct and so the slope on the right was correct. However, the left point is incorrect, and as a result the slope on the left side is incorrect, as in the x-intercept. When looking at the graph as a whole, the difference is slopes is not particularly noticeable, but if upon counting boxes, one would see that the slope on the right is -1/2, while the slope on the left is -1.

Graph 5 is a particularly interesting graph because precisely how much of it is correct can be interpreted in so many different ways. From one perspective, the slope on the right is correct, and the slope on the left is incorrect, so all the points on the left are incorrect, rendering the graph exactly half correct. From another perspective, 1/3 of the marked points are incorrect, rendering the graph 2/3 correct. From a regression perspective, the points on the solution are all fairly close to the points of the true line y=-(1/2)x-(1/4), rendering the graph almost entirely correct.

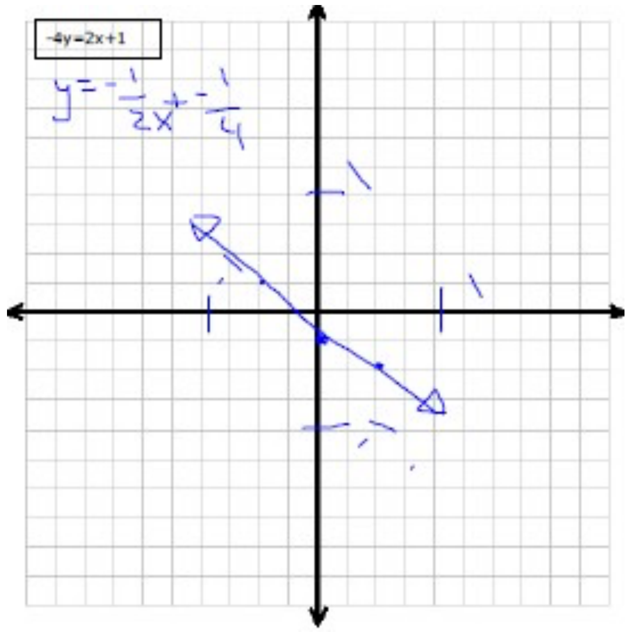

Figure 4. Student solution 5, for -4y=2x+1

The graders used none of the above justifications for their grades, but rather used a y-intercept and slope perspective in grading. The graders graded Graph 5 as mostly correct. Prior to discussion the mean grade for this graph was 7.73.

Only Michael mentioned the algebra mistake of writing x below the fraction line and took off 1 point. Michael gave 4 points for this graph because of the correct y-intercept, and he initially thought the slope was totally wrong. After the interviewer discussed with him the two dark points on the right side of the graph he decided the slope was correct. The interviewer then reminded him that the x-intercept was incorrect by going up the line on the left side. At the end of discussion he settled on 6 points for the graph.

The student set 4 small squares as a unit 1, but Daniel did not realize this when he graded the graph and he gave 5 points for the graph only, and after discussion he changed the grade to 8.

Sam took off 2 points, not because of the incorrect x-intercept, or point on the left, but because of the lack of labeling.

Frank, Gerry, Dolly and Alice gave full credit for this graph, and they all had the grading standard of 5 points for the correct slope and 5 points for the correct y- intercept. It was clear that they only looked at the close point on the right side to determine that the slope was correct. After the interviewer identified that the marked point to the left was incorrect, Gerry and Alice changed the grade to 9 and 8 points, respectively. However, Frank and Dolly considered the correct y-intercept and a single close-by point were good enough to show that the student understood how to graph a line.

Ray, Jimmy and Kara checked the x-intercept all the time. They judged the x-intercept incorrect and took off 2 to 4 points. Sandra judged that the line was curved up and so the graph did not have a correct slope, and she took off 3 points. When Alice, Daniel, Dolly, Frank, Gerry, Michael, Sandra and Sam were grading the graph, they did not look at and talk about the left side of the graph. Therefore, most graders only looked at the y-intercept and the other point on the right side before discussion.

After discussion the mean grade for this graph was 7.91. Only 4 out of 11 graders took off any points based on the left side of the graph. And 3 of these 4 graders looked at the left side because the x-intercept was on the left side. If the x-intercept had also been on the right side of the graph, we could not be sure whether they would have graded the left side at all.

This result is particularly interesting because it appears to be strongly culturally biased. If the graders had been culturally inclined to attend to points to the left and neglect points to the right, the scores would have presumably been very different.

Reluctance to Model Student Reasoning

When a student a graph judged incorrect it may be that the student has made errors prior to the graphing that resulted in an incorrect graph despite having the skill, for example, in graphing lines, students frequently make algebra mistakes that result in drawing the wrong graph. Two graphs were particularly useful for studying the grader's willingness to model the origin of a student's graphing error.

In the graph F (Figure 2), the student did not show any algebra work. However, if we examine the line graphed, we find that the student drew a line of the equation y= x+1. It was possible that the student made a mistake when he or she simplified the equation, such as solving for y as y=x-2+3=x+1. If this was the case, then the student did have skill to find points on a line and draw a graph from those points, and he or she only had an error in the manipulation of the equation prior to finding those points.

In graph 9 (Figure 5), all of the points are wrong and the slope is incorrect. However, if we compare the line in the graph and the correct line for the equation, they are reflections of each other across the x-axis. Therefore, it may be that this graph resulted from a sign error, slope and intercept are positive when they should be negative.

In the case of Graph F (Figure 2), Sam hypothesized about this student’s reasoning in creating the graph when he was grading. Sam gave the student 7 points. The other graders only gave the points for the correct slope – a feature of the appearance of the graph. Daniel said, “I will only look at what the students wrote or drew on the graph." After the interviewer suggested a possible reason why the student drew this line, none of the graders gave more points for this student.

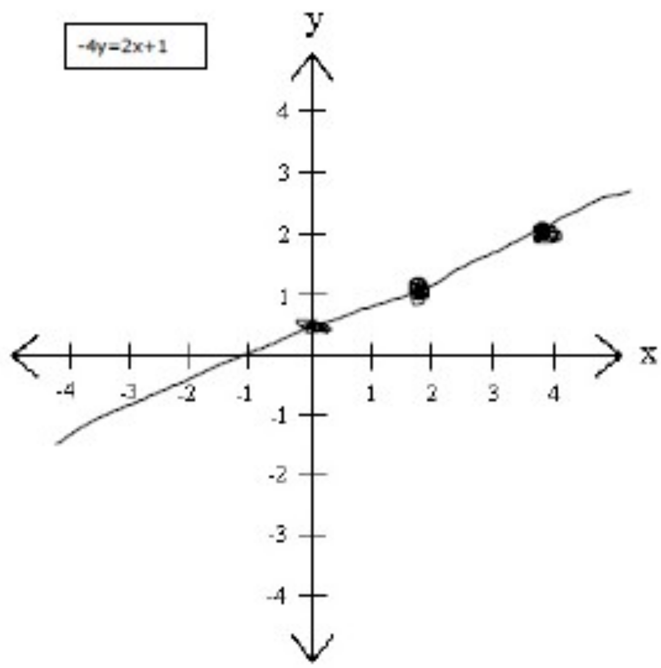

Figure 5. Student solution 9, for -4y=2x+1

The mean grades for graph 9 (Figure 5) were 1.36 before discussion. Only Kara, Ray, and Gerry judged the graph resulted from a sign error; Kara and Ray gave 3 points while Gerry gave 6 points for the graph. Gerry said, "I think the student put the negative sign on the right side when he simplified the equation. And the student knew how to graph so that I gave him 6 points." The other graders gave the graph 0 or 1 points. After each grader gave their grade, the interviewer discussed the possible reasons why the student created this particular graph. Only Alice and Sam changed their grades, both from 0 to 1. After the discussion, the mean grade was 1.55. Most graders graded based on the incorrect slope and intercepts shown on the graph and did not give any points based on the possible reasoning behind the graph.

From the graph F and graph 9, we see that most graders only graded the work shown on the graph itself and did not venture into the realm of modeling the students' reasons for creating the particular graph that they did.

Discussion

Altogether, the general grading rule appears to be: slope is worth 4 points, y-intercept is worth 4 points, labeling of intercepts, points and graph is worth 1 point. After that, add 1 point if everything else (such as curvature) is correct. Sandra said, "I gave 4 points for a correct slope, 4 points for a correct y-intercept, and 2 points for the graph as whole being correct." Her grading standard is very close to the grading rule summarized. Also, Sandra is the only grader who never changed any grade after the discussion. This reflects that this grading rule was very stable for her and applied to all kinds of graphs. Except for Sandra, this general rule does not express each individual grader's grading accurately.

There were some disagreements in the relative importance graders assigned to each feature, but the general rule is relatively close. All graders considered slope and y-intercept to be very important. Only some of them considered labeling to be important. Anything else was a matter of a single point adjustment. For example: in addition to the slope, y-intercept and labeling, some of the graders also scored based on the algebra, quality of the sketch or some other detail of the graph. They subtracted one point for each thing they noticed that was incorrect or added 1 point if everything seemed correct to them. Other graders only graded based on slope, y-intercept and labeling.

We also know the graders graded depending on the work shown on the graph and did not venture to think about the student's reasons for creating graph as they did. Instead, they graded almost entirely from y-intercept, slope, and labeling. More strikingly, graders paid attention to the right side of the graph only.

All together, this means that grading the slope doesn't really mean judging the overall steepness of the line or a relationship between aribitrary points. Instead, graders judged the slope from two specific points: the y-intercept and (usually) the first point to the right. If these two points were correct, then the graders judged the slope to be correct. Moreover, the graph was whole is then judged to be correct. Although the graders know a graph is more than two points, in this experiment they mostly only graded two points.

We suspect this style of grading is successful because graders have to grade a lot of problems at once, and two points is an efficient way to do it. In the math department studied, exams are pooled across sections. Each grader is assigned a page and grades the same problems across all sections. This means that the idiosycracies of an individual grader do not advantage or disadvantage their class compared to other sections. When grading a stack of problems, such as in this experiment, it is of interest to the grader to get through the stack as quickly as possible. Having only to features to look at and having those features always be in approximately the same place as long as the graph is correct makes judging a correct graph and scoring an incorrect graph much faster than alternative methods.

Due to the high consistency between interviwed graders that was not affected by education, and the high volume pooled grading system at this particular institution, we believe the two point rule to be representative of this department as a whole, and we suspect that it generalizes to similar situations of high volume grading.

Influence on Student Learning

This type of grading, with an emphasis on two points, may have an effect on student learning. In all the 20 used for the college algebra graders' interviews, 6 of them did not have any points plotted, 5 of them had points plotted on both left and right sides, none of them had points plotted only on the left side, and 9 of them had points plotted only on the right side. We can see the students had a tendency to only plot points on the right side of the graph. Although we cannot say for certain that feedback from graders over the student's career may play a large role in this, there does appear to be a culture among both graders and students of prioritizing the right side of the graph.

We are particularly struck by the similarity between the two-point grading style and Thompson's [13] case of "Sandra," where a high school math teacher had difficulty finding the y-intercept (and subsequent equation) of a line because of an excessive focus on slope as the rise and run between the pair of given points. That case suggests that not only does two-point perspective fail as students progress to more sophisticated mathematics; it also fails immediately when students are asked to move from finding points from given equations to finding equations of lines from given pairs of points.

Conclusion

When grading large numbers of papers quickly, it seems that the most common and most efficient manner of hand grading graphs of lines is to look only at the y-intercept and the first (convenient) point to the right. There is a danger that – as a result of this type of grading -- students understand graphing a line as just plotting two points and then connecting them. Beginning at the y intercept, students then go "over one and up m" to graph a line with slope m, or decomposing a fraction into "rise over run." In both cases, a slope is thought of as being composed of two discrete points.

This method is good enough to generate graphs of lines, from equations, but begins to fail as students and teachers work to find equations of lines from points [13]. Fundamentally, math teachers have a responsibility to not just teach to the problem immediately in front of them, but also take into account the prerequisite structure of mathematics, and prepare students for the math that is to come. Encouraging students to focus on points also does nothing to remedy the problems found in the literature, where students struggled to reason about intervals instead of points [3,4,6,8,9]. Focus on regularly points also does not prepare students well for higher mathematics such as elementary differential equations or dynamical systems [2,4].

For teaching graphs, there are several alternatives. Bell and Janvier [6] recommend that students be asked to identify the type of variation before beginning to create a table or plot points. Thompson [12] recommends a graphing activity in which students track each variable's changes in time individually before combining them into a graph, and has developed an entire framework for studying student's reasoning about dynamical systems based on this approach [14]. Paoletti et. al. [11] have further developed this approach into a local instruction theory.

However, these recommendations address instruction, not grading. Grading is less well researched, and we invite discussion on how grading might be improved to emphasize intervals rather than a small number of points. This brings us back to the original question of the study: Can we replicate human grading by machine? Yes. We only have to teach the computer to identify the (roughly two) points a teacher would look at to grade from and calculate the slope. However, computer grading also opens up new possibilities. The two-point grading method seems to exist for efficiency. If a computer is doing the grading, the computer can look at more than just those two points without making more work for the teacher. Alternative systems, such as linear regression, may reduce teacher workload while emphasizing all points equally. Such a change in grading method would not improve instruction by itself, but rather reduce the workload on teachers and change the method of assessment in such a way as to open the door to teaching and learning superior conceptions of line and slope a little wider.

References

1. Ye, Xiaojin, (2011). *Grading criteria of college algebra teachers.* Unpublished master's thesis, Kansas State University, Manhattan, KS.

2. Castillo-Garsow, C. (2010). *Teaching the Verhulst model: A teaching experiment in covariational reasoning and exponential growth*. Unpublished doctoral dissertation. Arizona State University.

3. Castillo-Garsow, C. (2012). Continuous quantitative reasoning. *Quantitative reasoning and mathematical modeling: A driver for STEM integrated education and teaching in context*, 2, 55-73.

4. Castillo-Garsow, C. (2013). The role of multiple modeling perspectives in students' learning of exponential growth. *Mathematical Biosciences & Engineering*, 10(5&6), 1437-1453.

5. Castillo-Garsow, C., & Castillo-Chavez, C. (2015). *Why REUs matter (p. 125).* New Jersey: World Scientific Publishing.

6. Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the learning of mathematics*, 2(1), 34-42.

7. Clement, J. (2000) Analysis of clinical interviews: Foundations and model viability. In Lesh, R. and Kelly, A.*, Handbook of research methodologies for science and mathematics education (pp. 341-385)*. Hillsdale, NJ: Lawrence Erlbaum.

8. Kerslake, D. (1977). The understanding of graphs. *Mathematics in School*, 6(2):22–25.

9. Leinhardt, G., Zaslavsky, O., and Stela, M. K. (1990). Functions, graphs, and graphing: tasks, learning, and teaching. *Review of Educational Research*, 60(1):1–64.

10. May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, *261*(5560), 459-467.

11. Paoletti, T., Gantt, A. L., & Corven, J. (2023). A Local Instruction Theory for Emergent Graphical Shape Thinking: A Middle School Case Study. *Journal for Research in Mathematics Education*, *54*(3), 202-224.

12. Thompson, P. W. (2002). Didactic objects and didactic models in radical constructivism. In *Symbolizing, modeling and tool use in mathematics education* (pp. 197-220). Dordrecht: Springer Netherlands.

13. Thompson, P. W. (2013). In the absence of meaning. In Leatham, K., editor, *Vital directions for mathematics education research*. Springer, New York.

14. Thompson, P. W., & Carlson, M. P. (2017). Variation, covariation, and functions: Foundational ways of thinking mathematically. *Compendium for research in mathematics education*, *421*.