Line of "Good" Fit in Grade 8 Classrooms

Annelise Nielsen

David Yopp

Anne Adams

Context for The Research

- Longitudinal Learning of Viable Argumentation in Mathematics for Adolescents (LLAMA)
  - 4 year study
  - 32 participating teachers
  - Argumentation intervention

Grounded Theory

- Developed materials without reference to existing literature
- Multiple phases of instruction and data collection
- Research Questions developed during phases of teaching experiment

Three Previous Studies

- Sorto, M. A., White, A., and Lesser, L.M. (2011). Understanding student attempts to find a line of fit. *Teaching Statistics 33*(2), 49-52.
- Casey, S. A. (2015). Examining student conceptions of covariation: A focus on the line of best fit. *Journal of Statistics Education*, *23*(1), 1-33.
- Anna E. Bargagliotti & Celia Rousseau Anderson (2017) Using Learning Trajectories for Teacher Learning to Structure Professional Development, Mathematical Thinking and Learning, 19:4, 237-259, DOI: 10.1080/10986065.2017.1365222

Previous Studies- Sorto, White, & Lesser 2011

- University Students
- What criteria do students use to draw a line of fit?
  - equal numbers of points on either side of the line;
  - in the middle of all points;
  - follow the data trend; and
  - distance of the points to the line.

Previous Studies- Casey 2015

- Grade 8 students
- What method do students use to place lines of fit?
  - passing through the most points;
  - dividing the points into an equal number on each side of the line;
  - connecting the first and last points;
  - and passing through the origin.
- How does this line compare to a least regression line?

Previous Studies- Bargagliotti & Anderson, 2017

- Teachers and Grade 8 Students
- What criteria do teachers use to critique lines of fit?
  - goes through the origin;
  - splits the points in half;
  - connect the first and last point;
  - and goes though the most points.

Research Questions

- What extant criteria do Grade 8 students use to choose the better line of fit between two lines “fit” to a set of data, when both lines express the trend of the data?
- Is a residual criterion accessible and useful to Grade 8 students when learning about line of fit?
- How does introducing a residual criterion impact student understanding of line of fit and their understanding mathematical modeling process?
- What stages of learning do students express as they engage in our lesson?

Existing Approaches in Grade 8

- “the line must follow the direction of the linear data well” and “ the line should have about half the points above the line and half the points below the line” (McCaw, et al. 2014)
- Close to most of the points (Math Connects Course 3)

Subjective Criteria

- Closeness
- Follows the trend

Objective Criteria

- Half the points above and half the points below

Residual Criteria

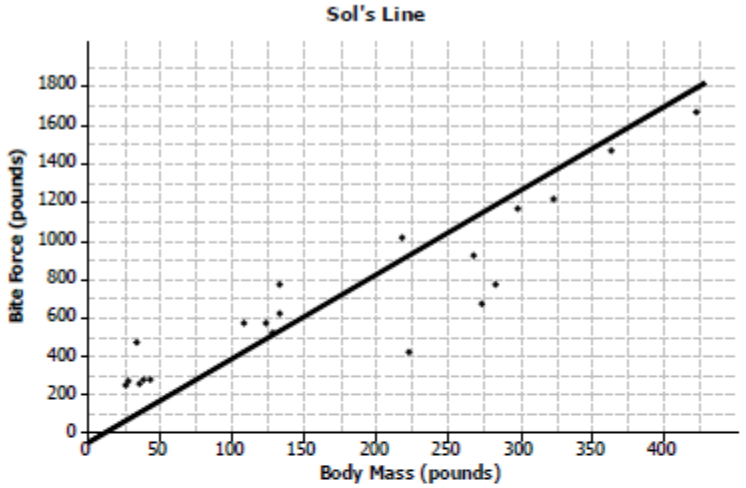
- Objective
- Required in high school CCSSIM
- The process of calculating residuals includes grade 8 content standard
- Hypothesized that it is accessible to grade 8 students

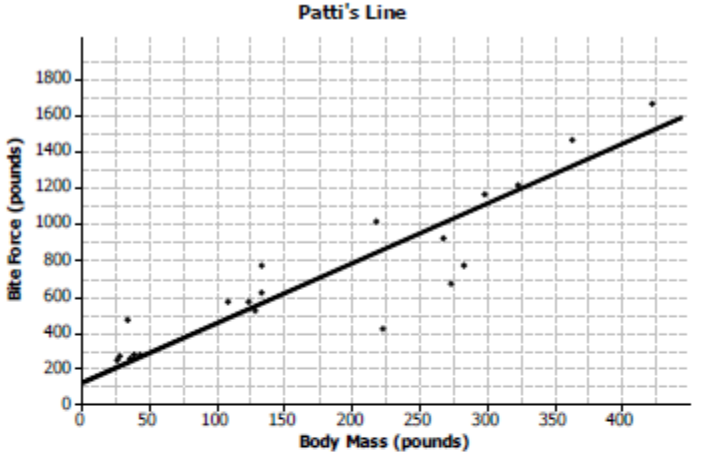
Phase 1

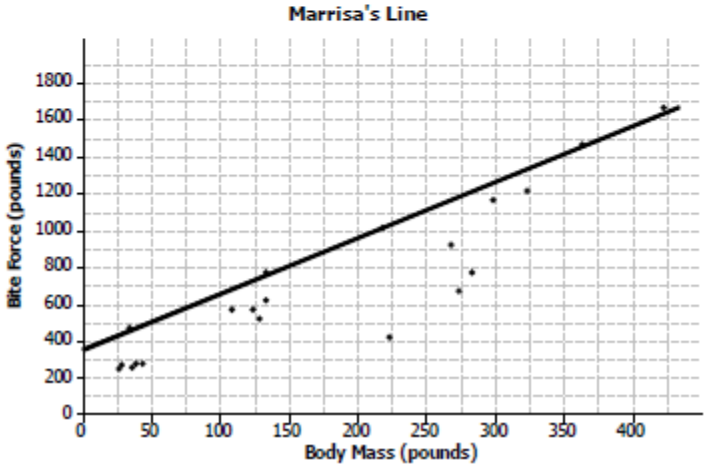
Anticipated progression

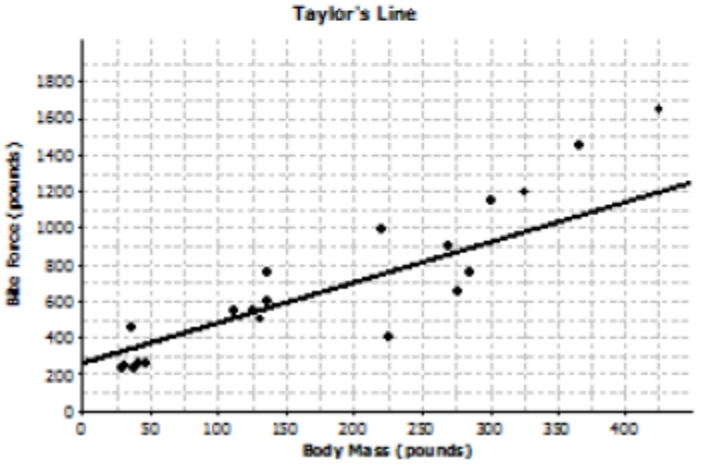
Part 1:

1. Students examine four lines of fit
2. Students develop a claim for each line stating whether it is a good fit for the data or not
3. Students develop an argument for their claim and explicitly state the criteria they used to make their choice.
4. Students list the criteria they used in each of the four arguments they developed.









https://www.engageny.org/resource/grade-8-mathematics-module-6-topic-c-lesson-9/file/48811

Phase 1 Continued

Part 2:

5. Students are introduced to the residual criterion
6. Students calculate the sum of the residual for two of the lines of fit
7. Students develop an argument for which of the two lines of fit is a better fit

Phase 1: Results

- Students used a wide variety of criteria
- Different students used the same criteria to support choosing a different line

Phase 2

- Consider only Pattie and Sol's lines of fit
- Added summarizing activity
  - Students share which line they believe is a better fit and why
  - Teachers record student responses
  - Teachers highlight instances when student's criteria are used to support different lines

Hypothesis: Summary activity will create disequilibrium and create the need for an objective criteria

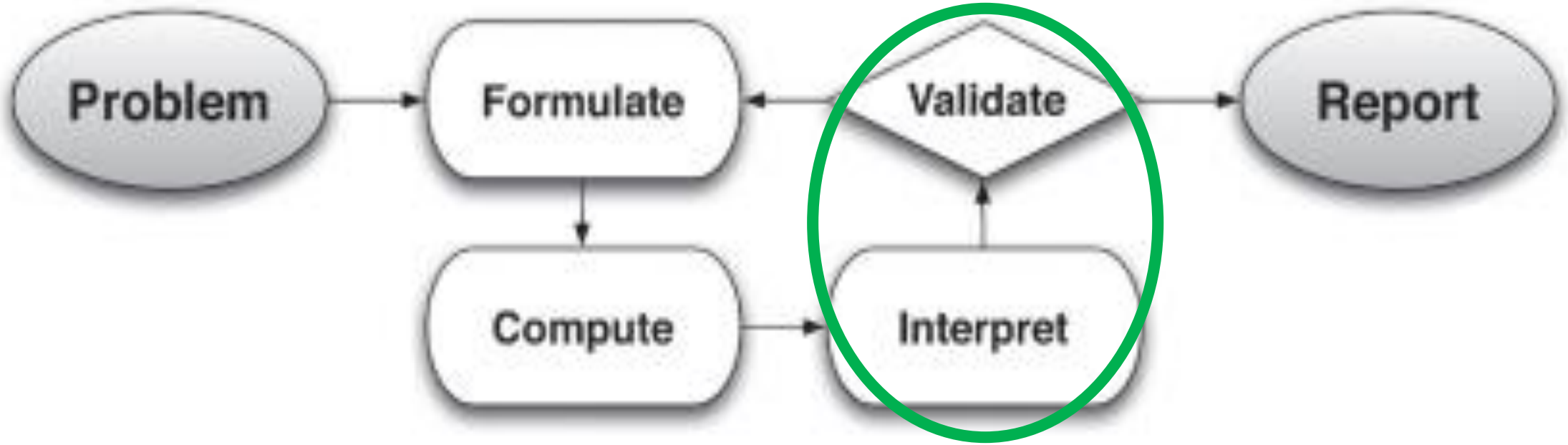
Phase 2: Results

- Some students use original criteria with residual criterion in their argument (Part 2)
- Many students use contextual criteria to make decisions about the lines
  - Should go through the origin
  - Does not makes sense for small crocodiles
  - Makes sense for large crocodiles

Phase 3

- No longer specify what criteria students should use for their argument
- Emphasized the modeling

Modeling Cycle



http://www.corestandards.org/wp-content/uploads/Math_Standards1.pdf

Phase 4: Final Progression

Part One

- Students examine two lines of fit for crocodile bite force data.
- Students develop a claim for which of the two lines is a better fit for the data.
- Students develop an argument for their claim and explicitly state the criteria they used to make their choice.
- Students share with the teacher/researchers the criteria they used, which are listed on the board by the teacher/researcher.
- Students acknowledge the variety of criteria used by their peers for determining the line of better fit, and they acknowledge that even when using the same criteria, peers came to different conclusions about which line is the better fit.

Phase 4: The Final Progression *Cont.*

Part Two

- Students receive explicit instruction of the least average residual (and least total residual) criterion for line of “better” fit.
- Students calculate the total and average residuals for each line by completing the following steps
  - Estimate the equation of each line of fit
  - For each model use the equation to calculate the estimated bite force for each crocodile (students are given the data table with the data set plotted)
  - For each model calculate the residual by taking the absolute value of the difference between measured bite force and estimated bite force
  - For each model sum the residuals and find the average residual
- Students develop a new claim and argument for which line is a better fit for the data.
- Students reflect on their extant criteria and the line chosen using these criteria to the least residual criterion and the line chosen with the least residual criterion.
- Students use their extant contextual criteria to restrict the domain of the line chosen using the least residual criterion to an interval that makes sense for the context.

Phase 4: Data Source

- 2 class periods of 1 LLAMA teacher
- Small city in Pacific Northwest
- 38.73% students eligible for free/reduced meals
- Students had already completed the planned instruction on line of fit

Findings: Grade 8 Extant Knowledge

The following criteria appeared across all implementations of the lesson

1) line splits points in about half above the line and half below,
2) the line crosses more data points
3) passes through origin
4) points closer to the line (vague)
5) more accurate (vague)

*Students in early iterations had not received instruction on line of fit and students in later iterations had already completed that unit of instruction

Part 1 Criteria Used by Students

	1st Class	2nd Class
Criterion Mentioned	Frequency Listed	Frequency Listed
Crosses more data points	1	12
Splits points in about half above the line and half below	9	5
Passes through the origin	9	5
More realistic to context (for small crocodiles)	2	3
Goes through the average of points (unclear)	0	2
One line is too steep/slope follows the trend in data	0	4
Points closer to line (vague)	1	2
Easy to calculate (vague)	1	0
More accurate (vague)	2	0
Easy to solve equation (vague)	1	0
Unclear	0	2
Number of students in the class	14	18

Part 2: Accessibility of The Residual Criterion

Categories	1st class	2nd class
Residual Criterion	11	16*
No Criteria	3	1
Easier to Work With	1	
Hits/Crosses More Points	3	1*
Starts at 0	1	
Vague	1	
Cannot have negative bite force	0	2*
Total Number of Students**	20	17

*Students who used multiple criteria
** Some students missed the first day of instruction due a school event

“Patties average residual was 131.75 and Sols was 168.25 showing that Patties average was closer than Sols. Based on the info shown [calculations of average residual for both lines] Patties equation is the better one out of the 2 and is the most accurate when it comes to how close the info was to the real info”

Findings: Impact of Residual Criterion

- Majority of students used the residual criterion to correctly choose the better line of fit
- 17 out of 27 interpreted residual as error

Findings: Impact of Residual Criterion- cont’d

- 6 out of 37 included additional criteria

“The residuals for [Pattie’s] line is smaller which means the actual and estimates were closer; also, you can’t have a negative bite force and a 0 lb crocodile.”

Findings: Role of Modeling Cycle

Origin Criterion

- Initial argument 19 of 37
- Final argument 4 out of 37

Discussion

- Instruction on line of fit inconsistent with Modeling Cycle
- ”Closeness” is problematic Criteria

Conclusion

- Students can leverage least average residual to select a line of fit
  - Potential to give access to the idea of “closeness”
- Residual criterion aligns with the modeling cycle
- Students can leverage context criteria to restrict the domain of their chosen model