Relative Difficulty of Physics Test Questions

G. Rieger

Physics & Astronomy

The Study

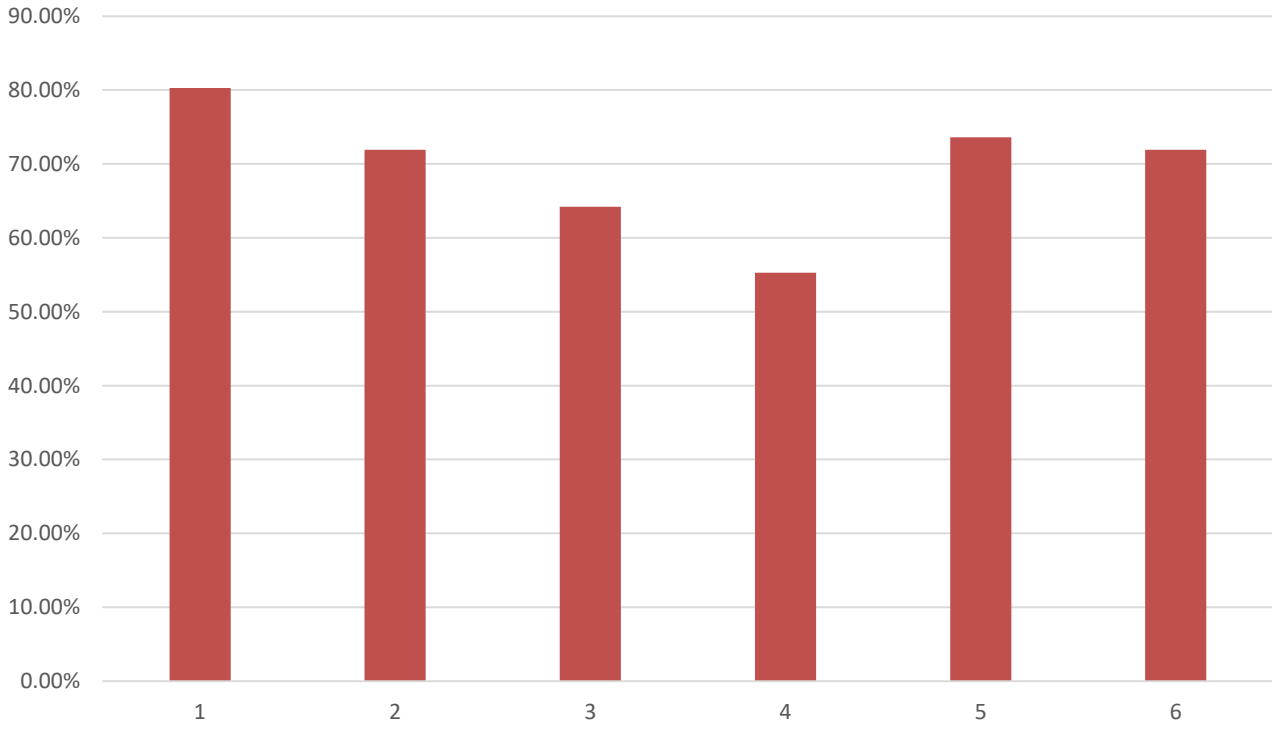
Physics questions from an online test bank are analyzed in terms of their difficulty level.

Each problem question comes in 5 – 6 similar versions (and randomized numbers) to reduce cheating.

Although intended to be similar, we generally find different outcomes for questions that are not exactly the same.

While not desirable from a fair testing point of view, it provides us with a large database that allows us to examine what **formal features** make a question more difficult.

Similar questions do not necessarily lead to similar outcomes.

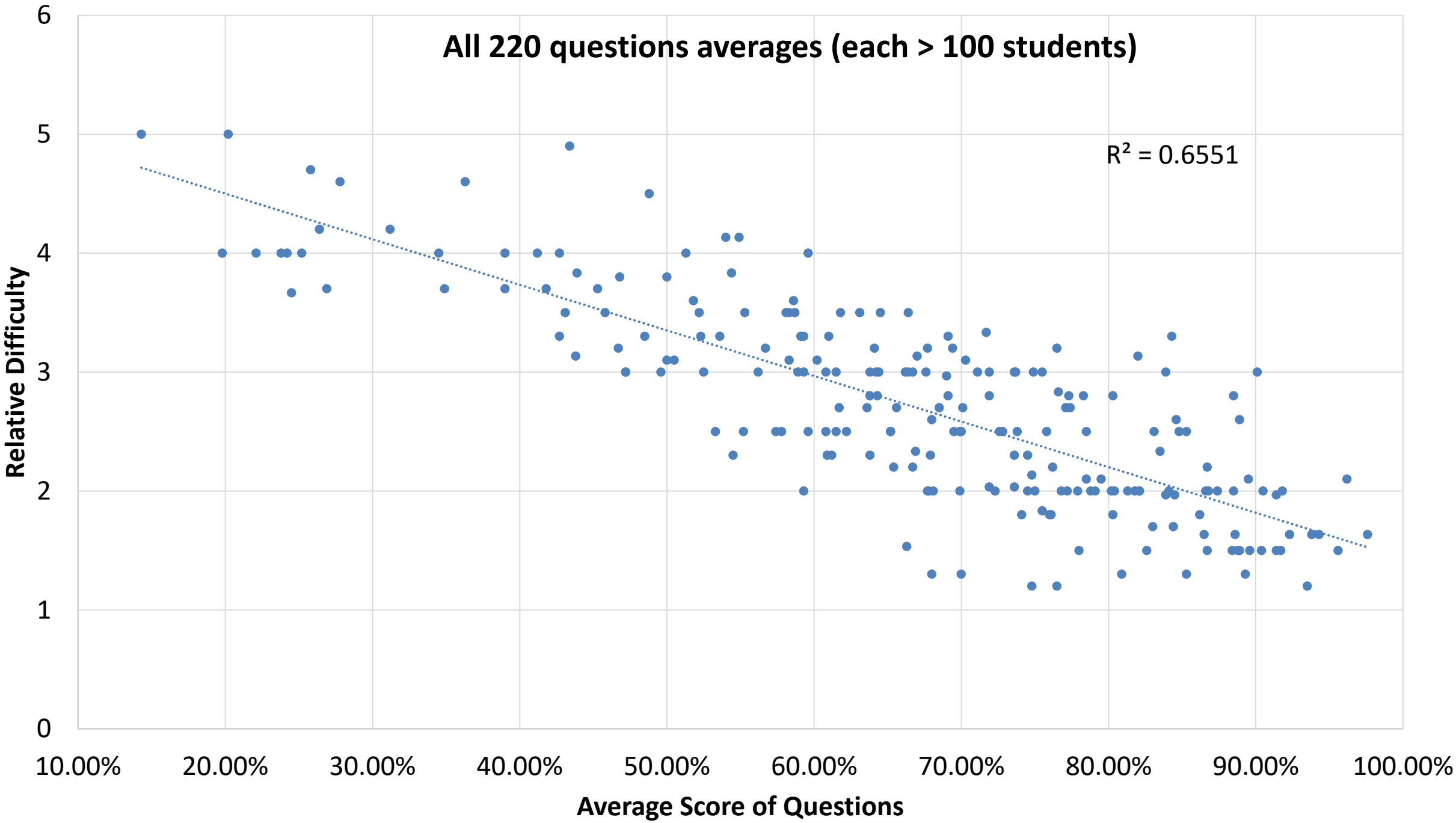


The graph shows six versions of a “person on bathroom scale in an elevator” question. The concept (and key equation) to solve this problem is exactly the same for all versions.

(1) and (2) are both mulitple-choice (MC) questions with four choices, but different quantities are given. (3) and (4) are numerical questions that have one additional step (max number of people or max acceleration). (5) and (6) are exactly the same MC question with six numerical choices and different numbers for the mass.

Using the scores, a common difficulty rating was generated.

Difficulty is here expressed as “difficulty points” that are derived from the formal question features.



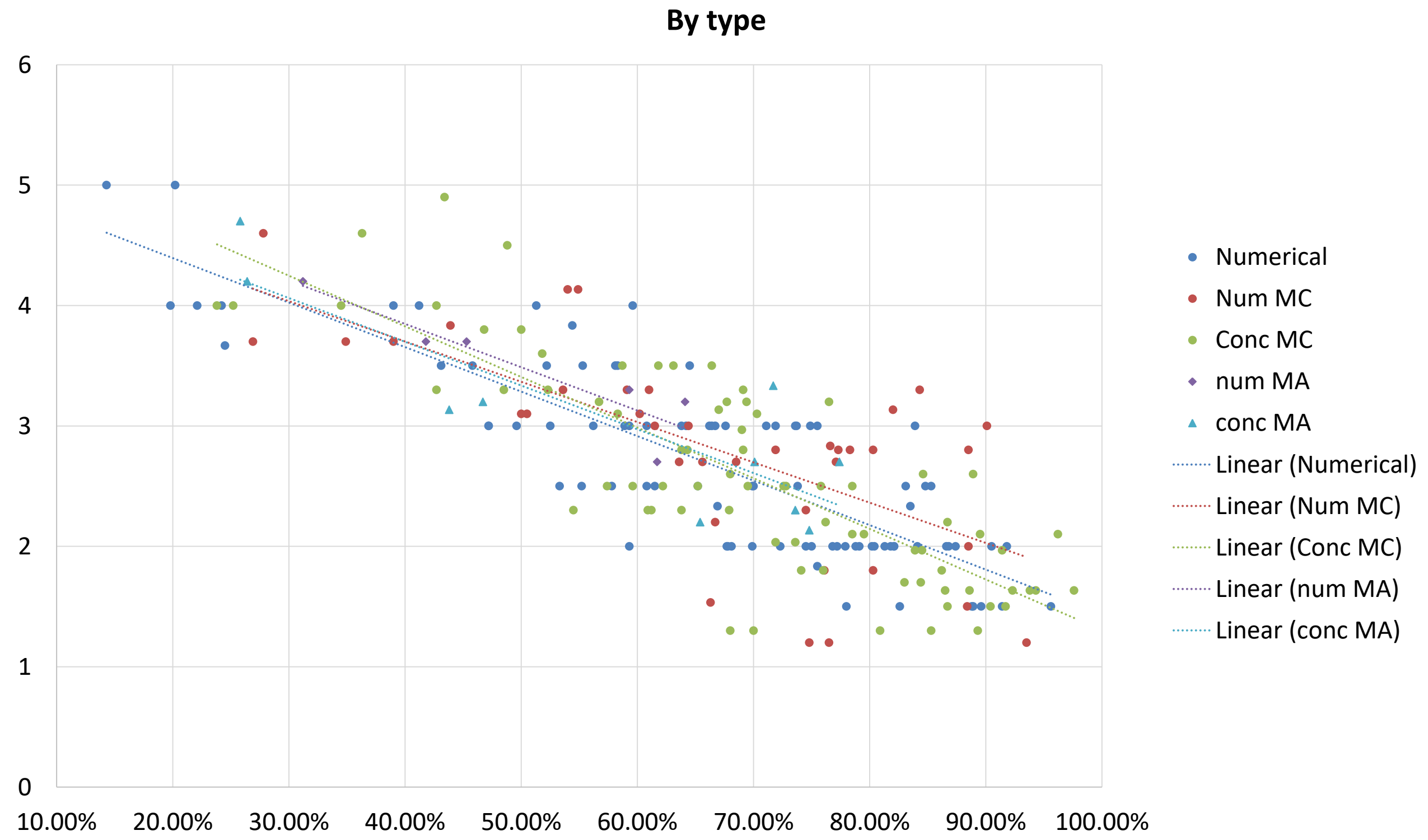
Difficulty points:

Numerical (open):	1
Numerical MC:	1 – 1/choices
Conceptual MC:	choices/5
Numerical MA:	choices/5 + correct/2
Conceptual MA:	choices/5 + correct/2
Solution Steps	+0.5 x Steps
Ratios/Relations answers	+0.5
Additional Concepts	+1
Info in Graphs	+0.33
Circuit Diagrams (n ≥ 2)	+0.5 + 0.5 x Junctions
Two Dimensions (angles)	+1
Known Misconceptions	+1
Unfamiliar Application	+1
Trick (wrong answer attractive)	+1
Extra Information	+1
Other Adjustments	+1 (Sqr/sqrt, hidden factor, unusual conversion, etc.)

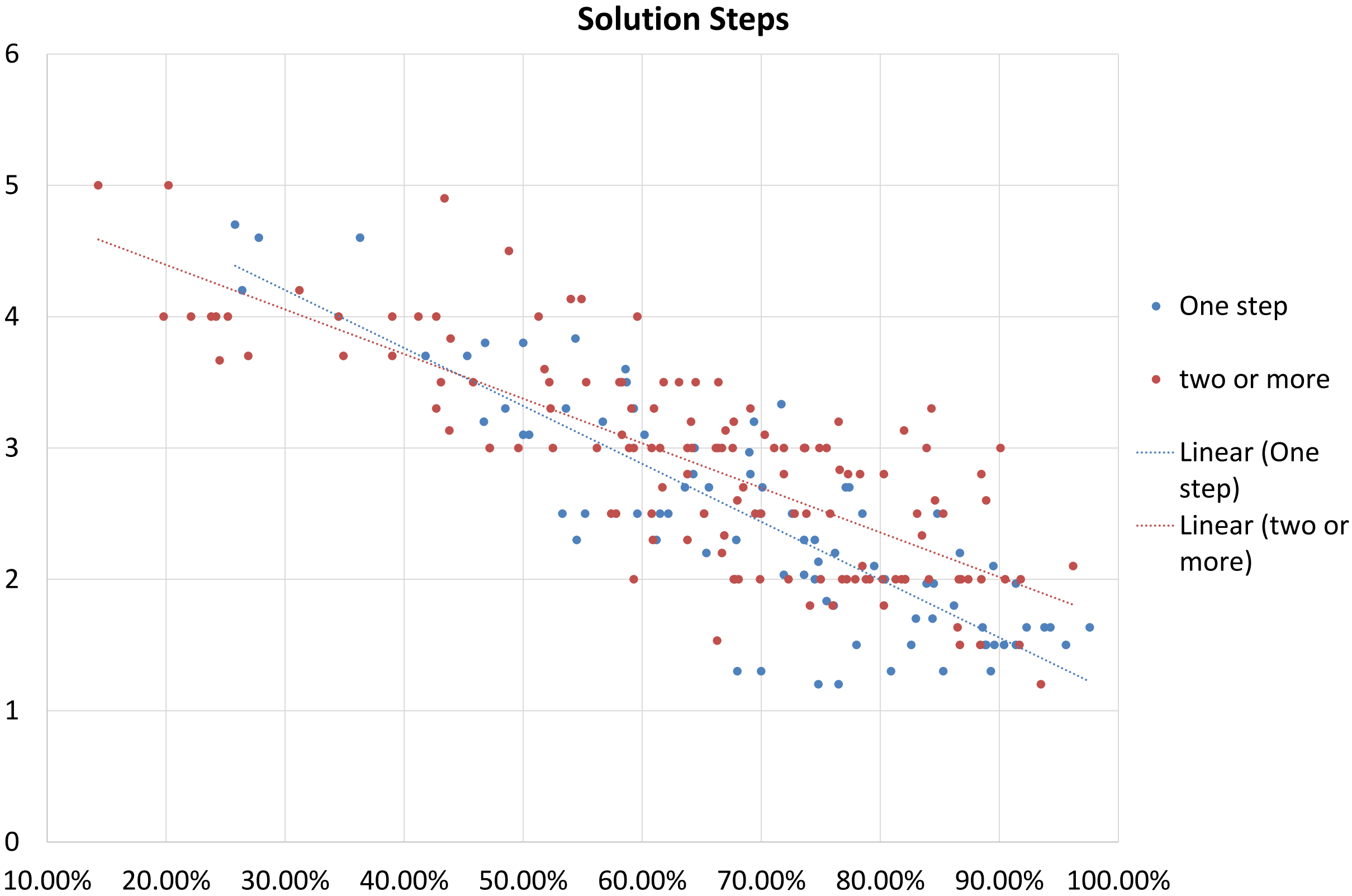
These are the base points depending on type

These are the points for extra difficulty

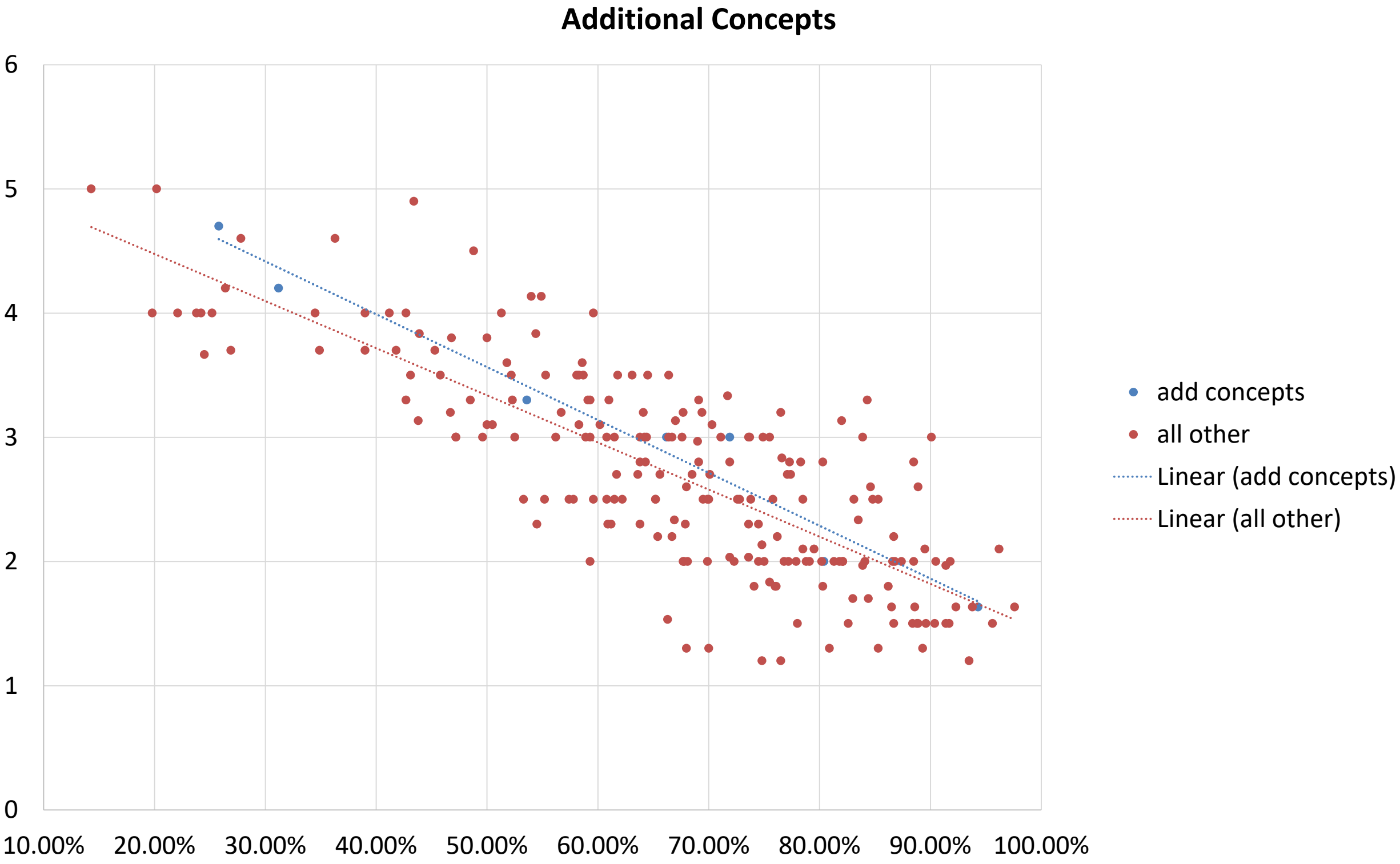
How well do the assigned points fit? Question type



How well do the extra points fit? Number of Steps

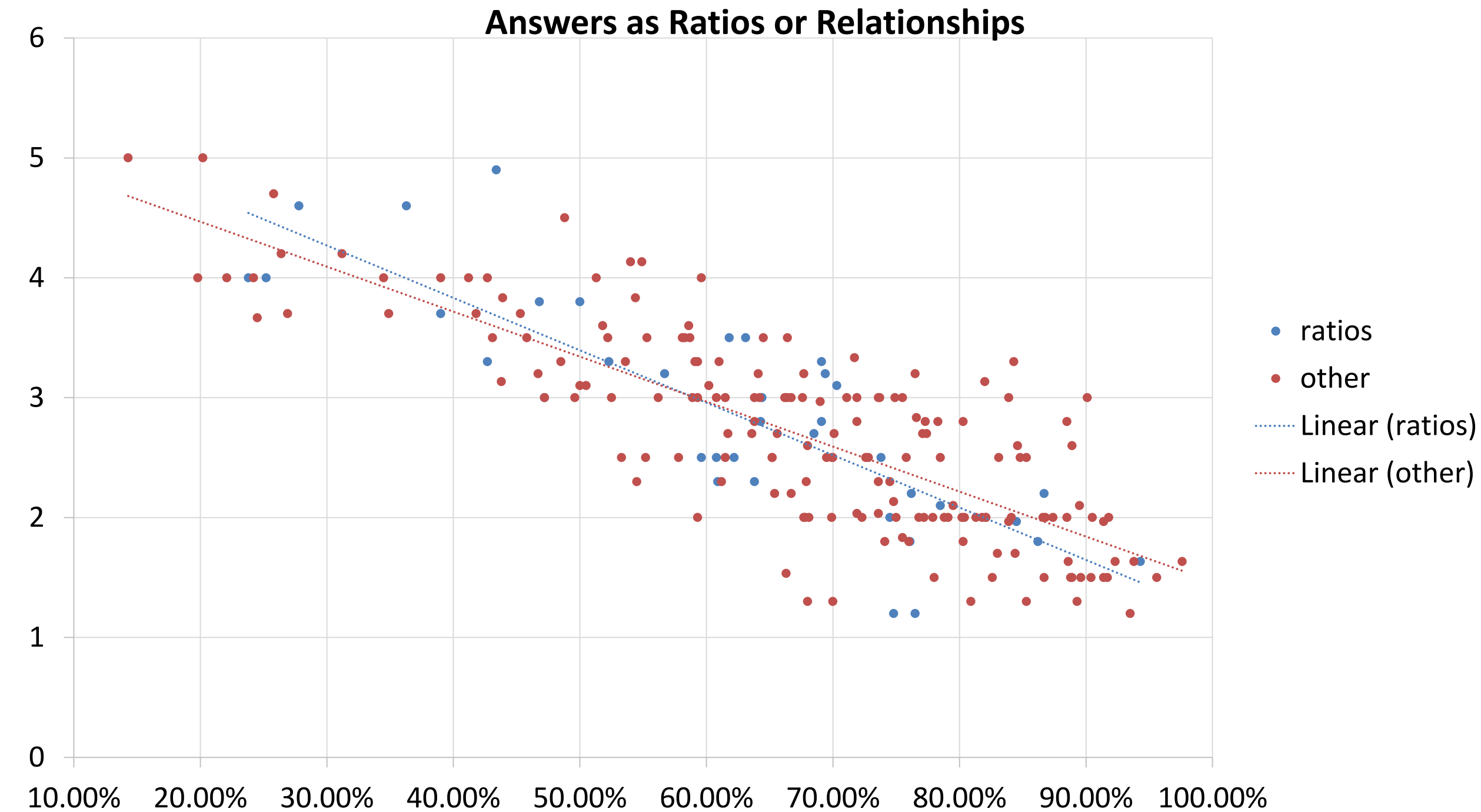


How well do the extra points fit? Additional Concepts



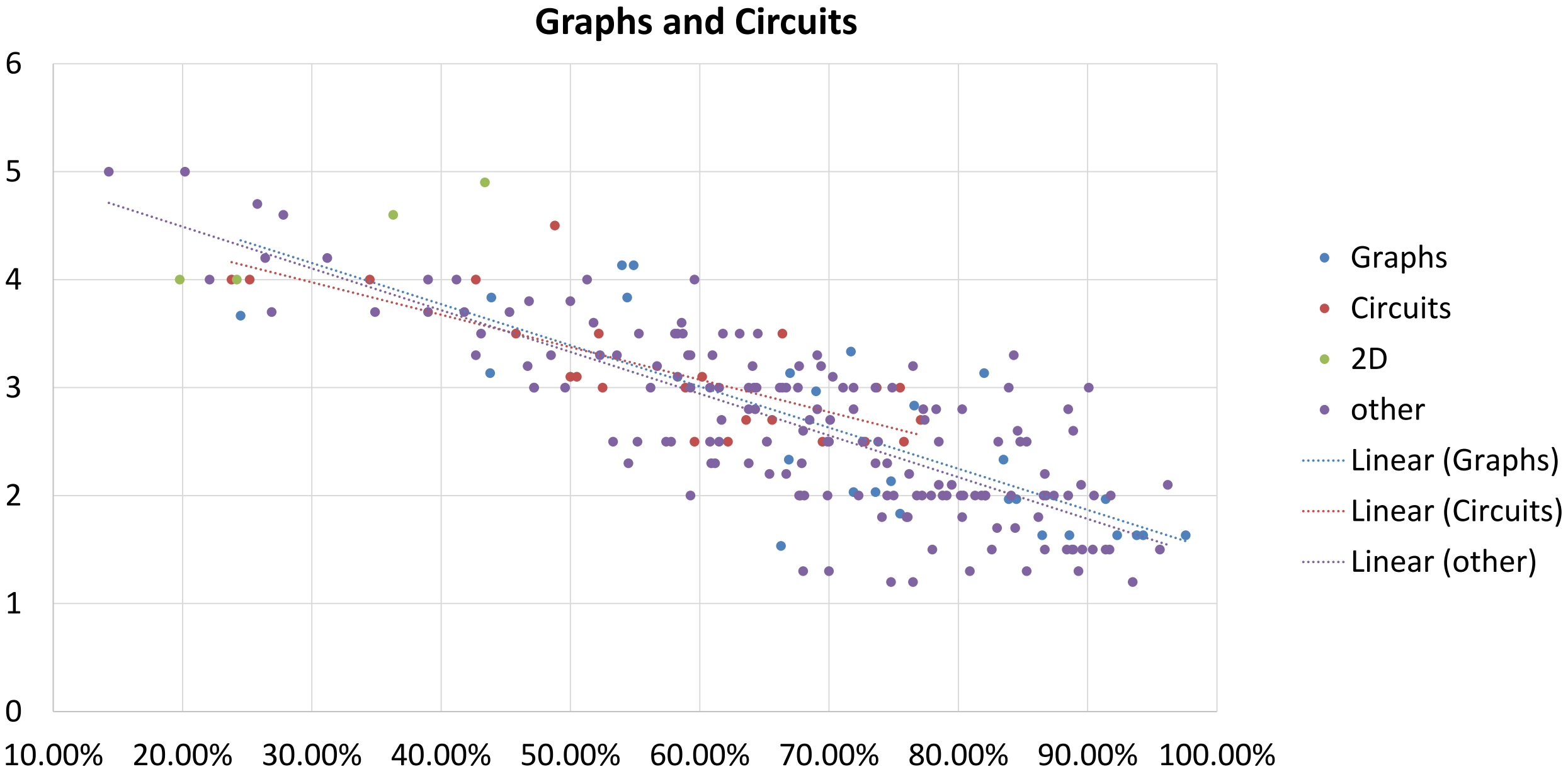
How well do the extra points fit?

MC/MA questions with ratios or relationships as answers.



Exception: key equation is one of the answers.

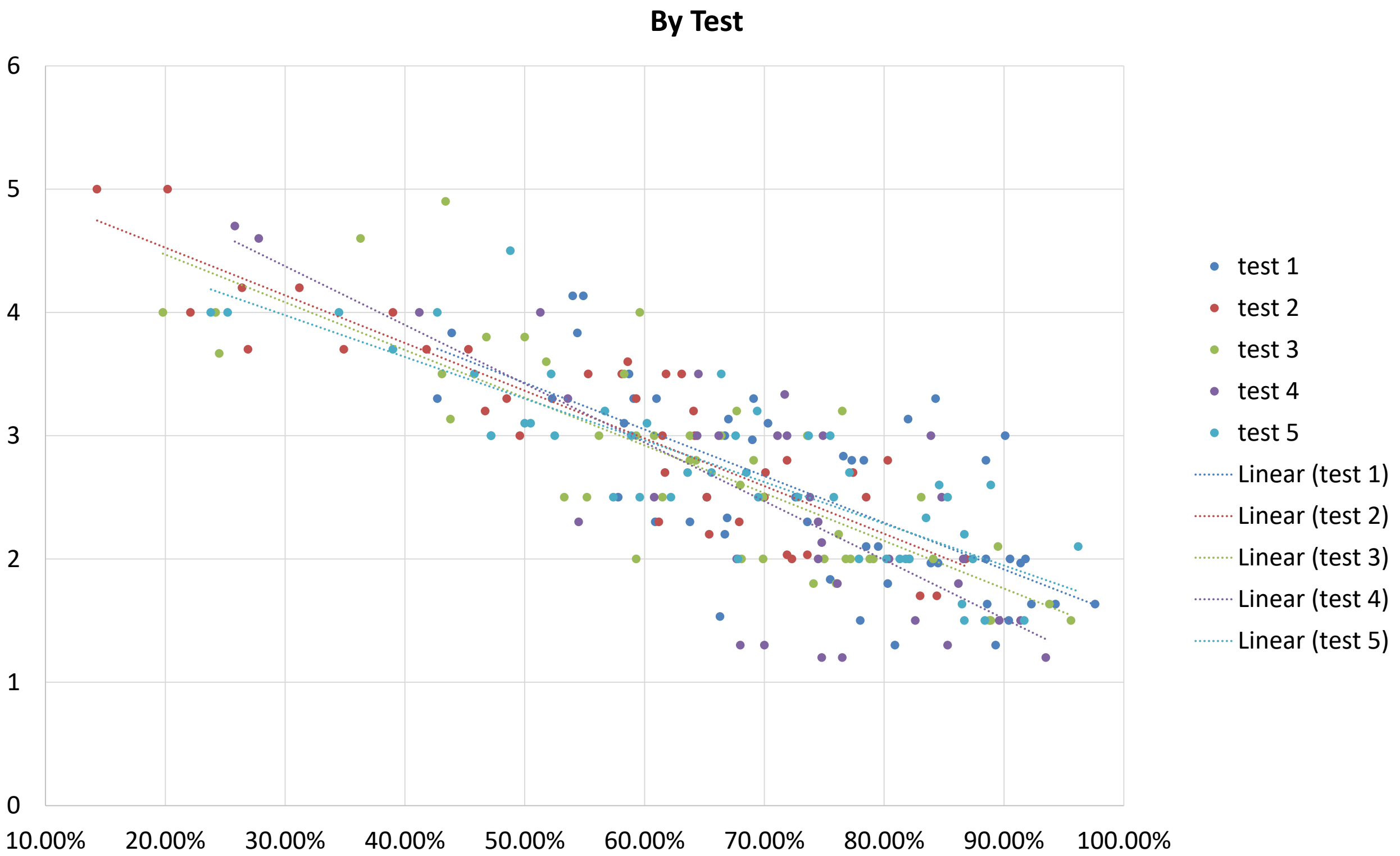
How well do the extra points fit? Graphs and Diagrams



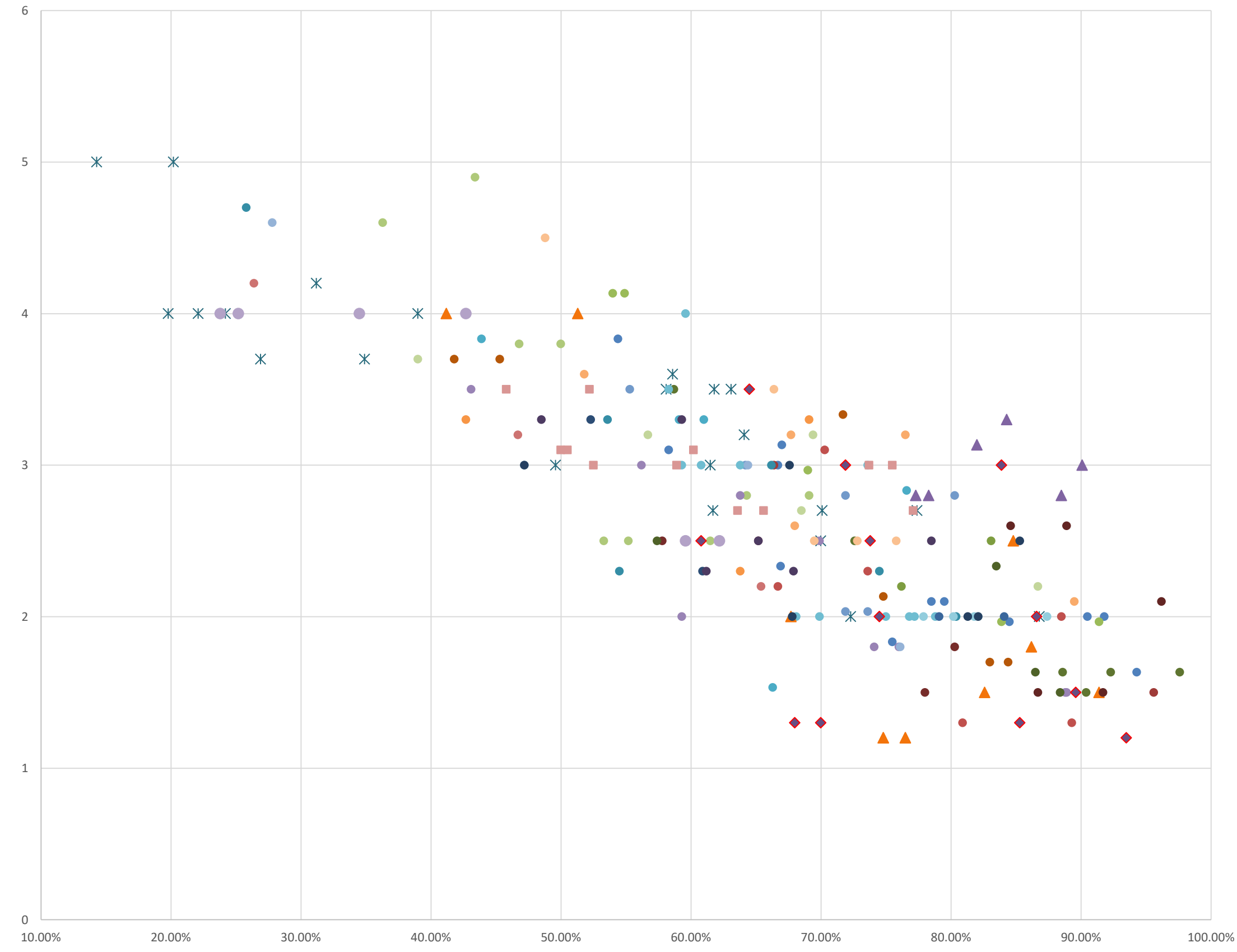
No extra difficulty due to FBDs or synchronous motion in x- and y-directions (e.g. hero jumping on moving train)

How do the tests compare?

1: Kinematics; 2: Forces; 3: Work and Energy; 4: Heat; 5: Circuits

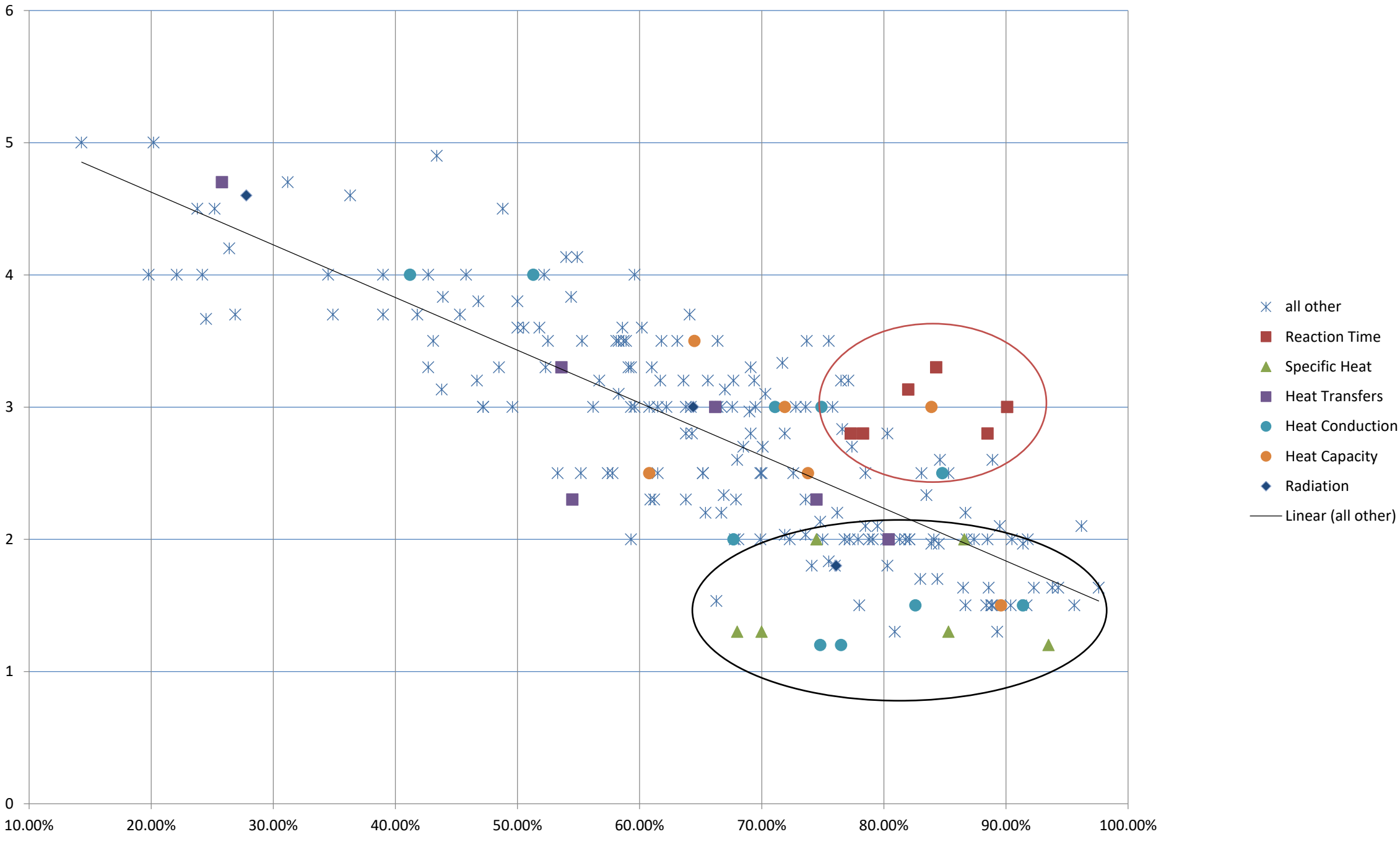


Topics



- Kinematics
- Free Fall 1
- Acceleration
- Reaction time
- Chase
- Free Fall 2
- Car braking
- Free fall 3
- Turning point
- 3rd law
- 2nd law
- Lab
- Elevator
- Interactions, FBD
- Work
- Energy, NC
- Efficiency
- Energy, cons
- Energy, cons
- kWh unit
- Power
- Fuel cons
- Specific Heat
- Heat transfer
- Heat conduction
- Radiation
- Resistor Circuits
- Resistance
- Bulb Circuits
- Electric Power
- Circuits, switches
- Current
- Units
- Lab 2

Topics: ‘Reaction time’ seems to be a well discussed topic – ‘Specific Heat’ not so much.



Final Exam Questions: Difficulty and Scores

y = -1.2505x + 3.4398
R² = 0.068

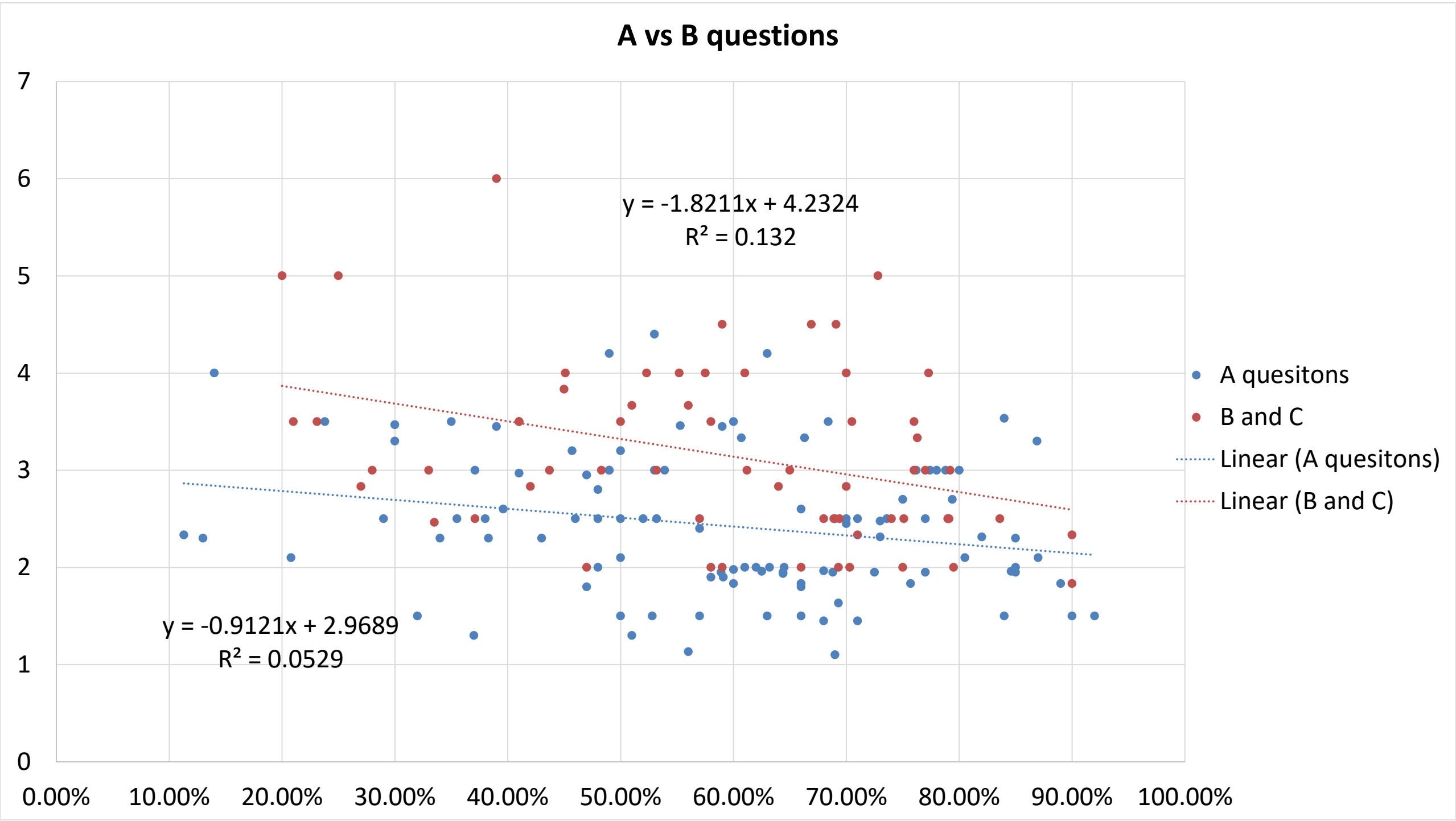
Relative Difficulty

7
6
5
4
3
2
1
0

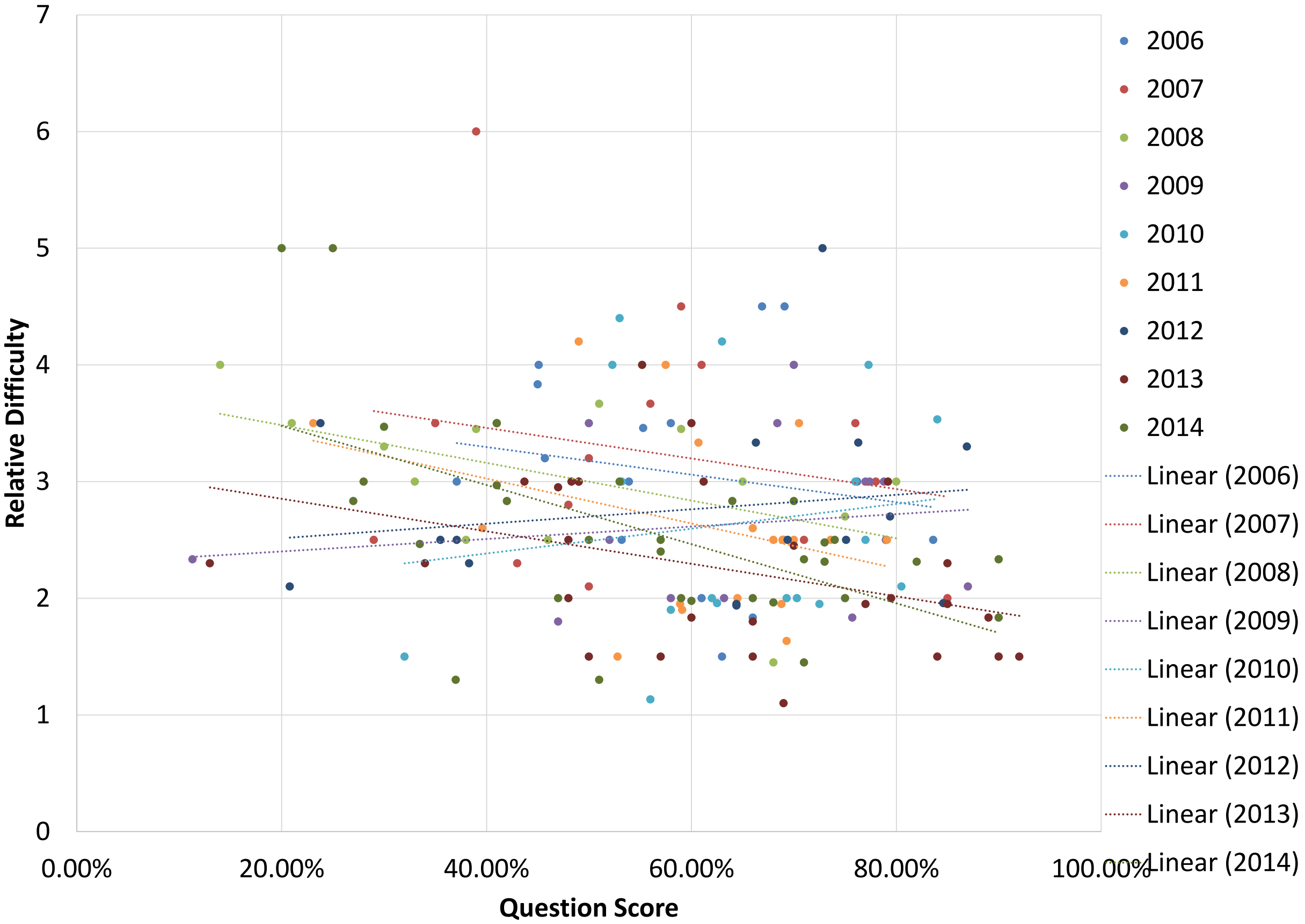
0.00% 10.00% 20.00% 30.00% 40.00% 50.00% 60.00% 70.00% 80.00% 90.00% 100.00%

Question Score

Final Exams – MC (A) vs Multi- Part Free Response (B and C)



Final Exams: Difficulty and Questions Scores



Conclusions

- Formal features contribute significantly to the difficulty of the test questions.
- General topic has no significant influence.
- The scheme allows identification of specific content that needs more support in class.
- The formal difficulty of questions does NOT have a significant influence on final exam scores.
- Hypothesis: final exam scores are mostly influenced by students ability to remember relevant concepts for each particular question. Once students are able to get started, they can usually solve a problem.