

# **The Tale of Two Tails:**

## **Adjustments to the scale**

**Hariharan Swaminathan**  
**H. Jane Rogers**  
**University of Connecticut**

# The Connecticut Vertical Scale

- **The Connecticut Vertical Scale (CTVS) places the mathematics and reading scores of children in Grades 3-8 on a continuum**
- **Designed to assess children's growth as they progress through the grades**
- **The vertical scale is a linear transformation of the Item Response Theory (IRT) proficiency scale**

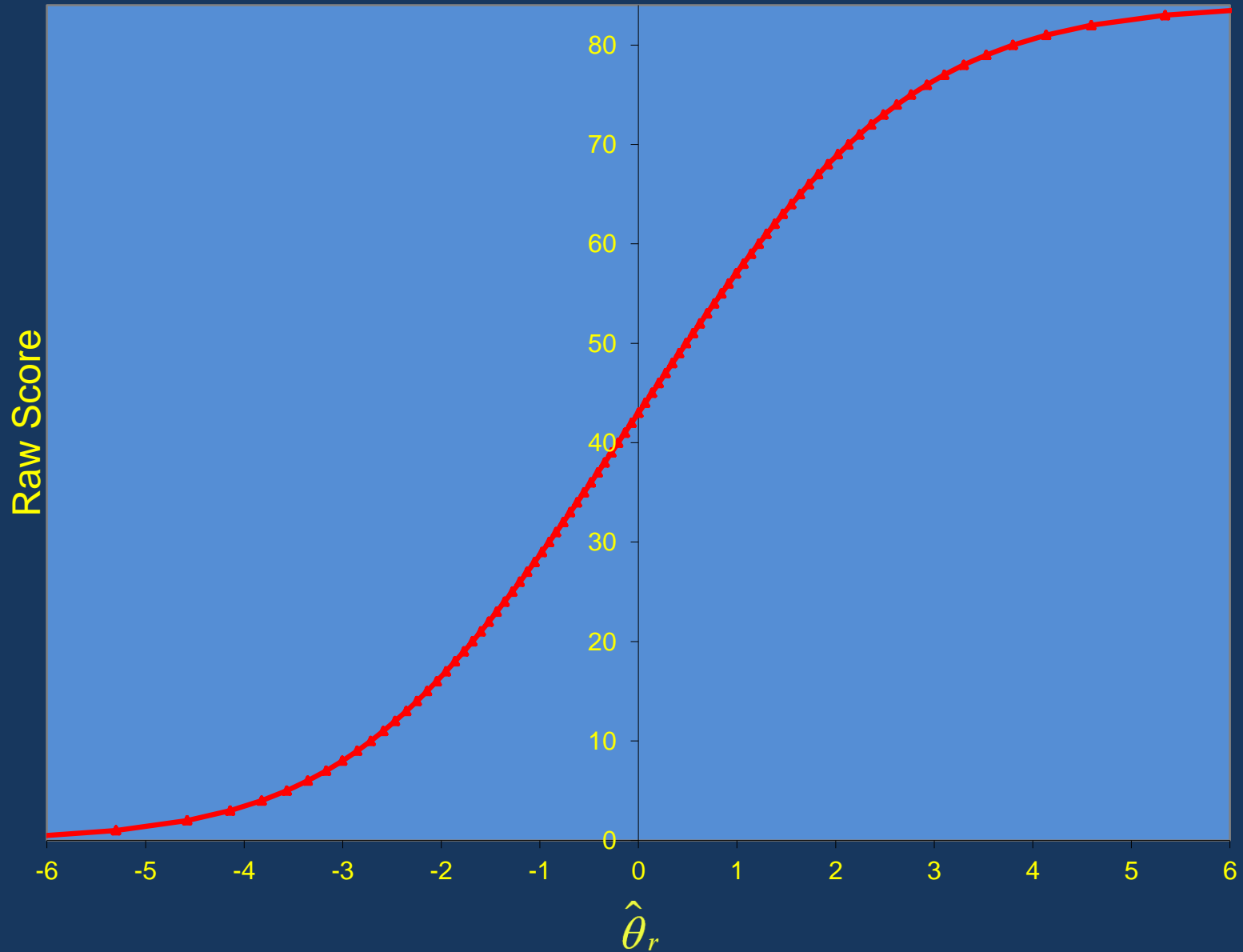
# Relationship Between Raw Score ( $r$ ) and Theta

$$r = \sum_{j=1}^n P_j(\theta_r)$$

$$P_j(\theta) = \frac{e^{(\theta - b_j)}}{1 + e^{(\theta - b_j)}}$$

$\theta_r$  -- MLE of Proficiency  $\theta_r$

Figure 1. Relationship Between Raw Score and  $\hat{\theta}_r$



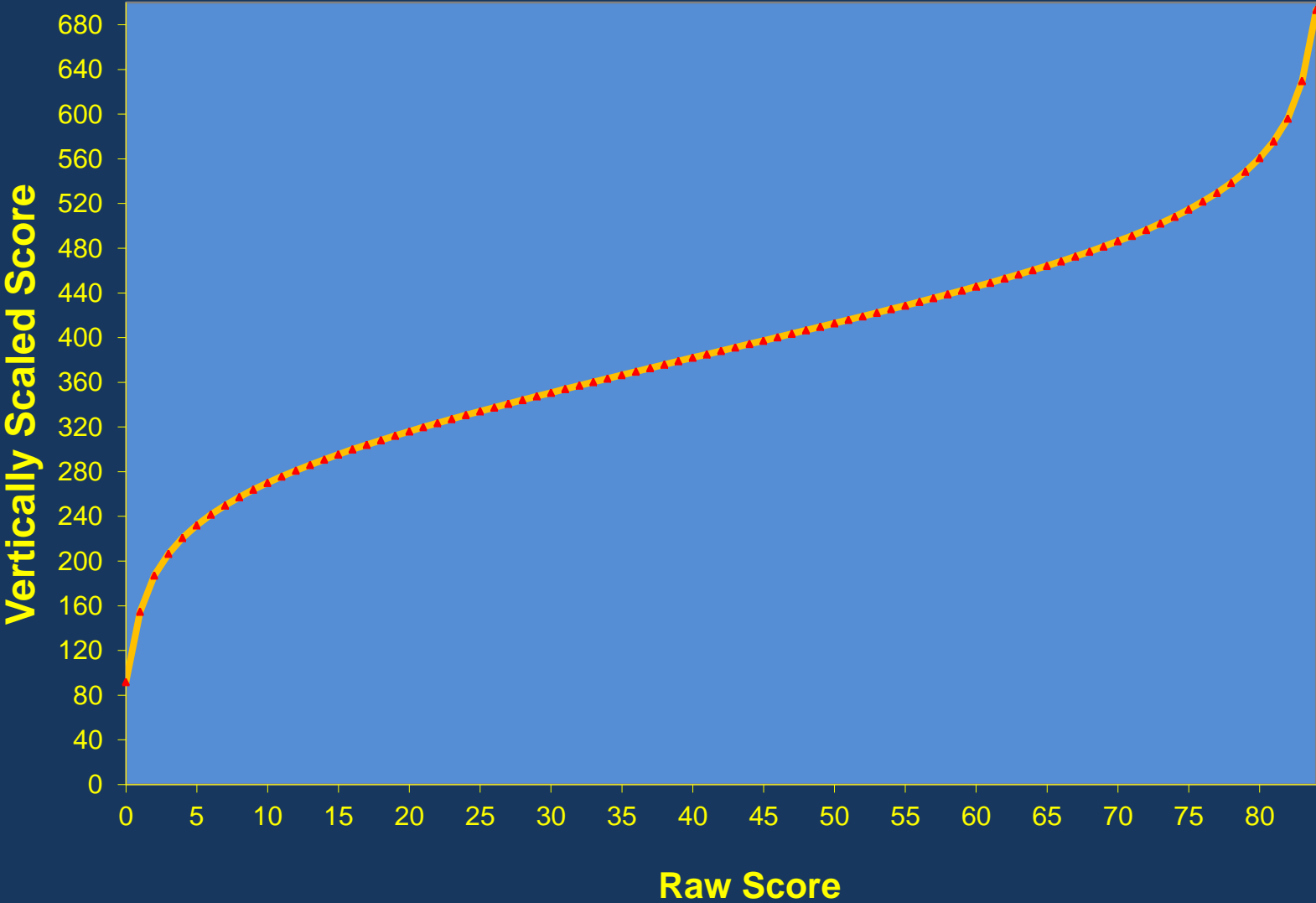
# Relationship of the Vertically Scaled Score (VSS) to $r$

The VSS is a linear transformation of  $\theta_r$

Hence

$$r = \sum_{j=1}^n P_j (VSS)$$

# Relationship Between Raw Score and Vertically Scaled Score

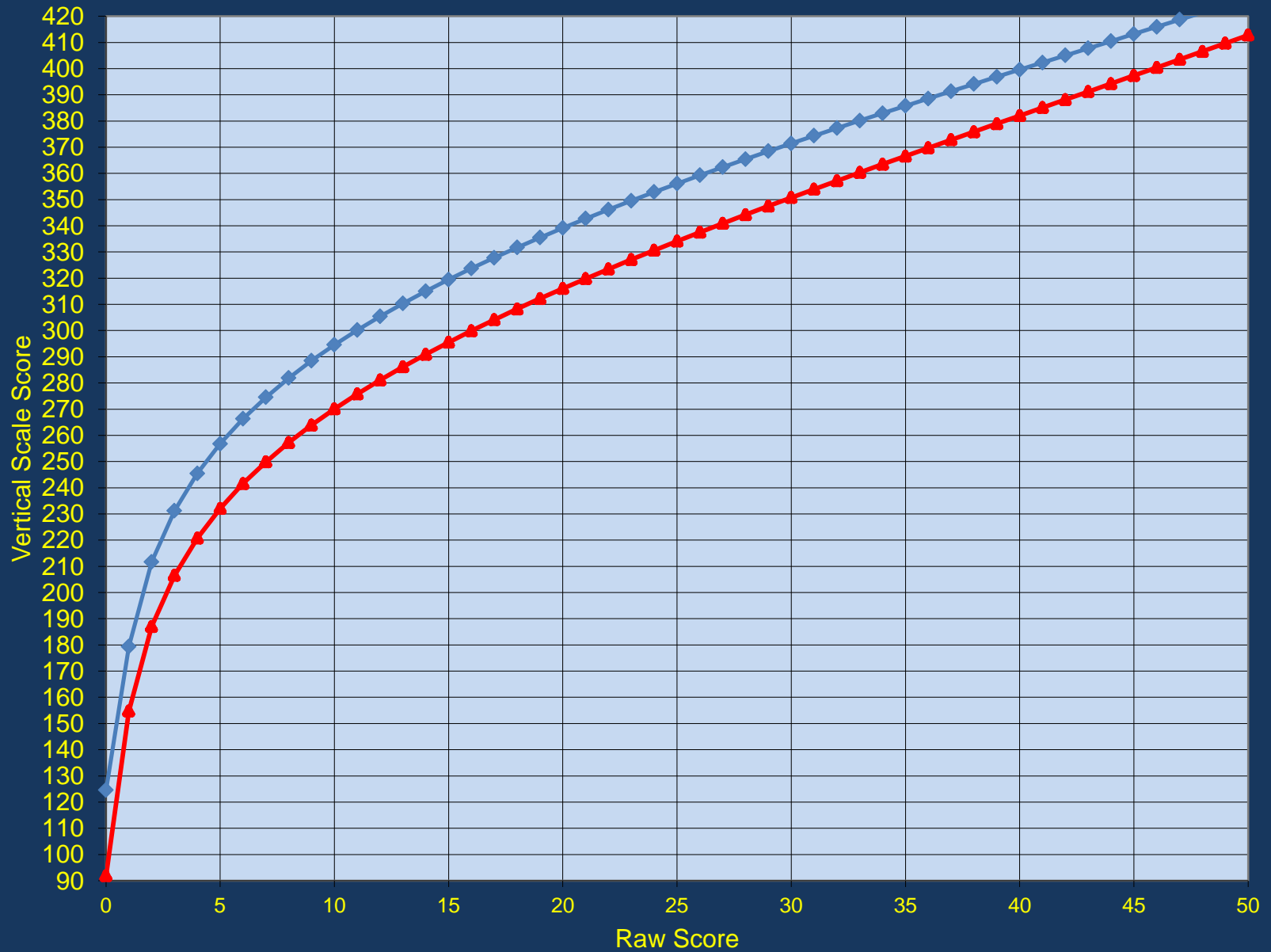


# Problem (?)

**Small changes in raw score at the extremes result in large changes in the Vertically Scaled Scores**

**This may cause concern to users when a child progresses from one grade to the next**

Figure 2. Relationship Between Vertical Scaled score and Raw Score For Children Progressing from Grade 3 to Grade 4 (Lower End - Reading)





**Figure 3. Relationship between Vertical Scaled Score and Raw Score For Children Progressing from Grade 3 to Grade 4 (Upper End - Reading)**



**Table 1. Gain in Unadjusted Vertical Scale Score Corresponding to an Increase of in Raw Score (LOWER END)**

Raw Score Grade 3	Gain in Unadjusted Vertical Scale Score Grade 4			
	Raw Score +1	Raw Score +2	Raw Score +3	Raw Score +4
0	87	120	139	154
1	57	76	91	102
2	44	58	70	79
3	39	50	60	68
4	36	46	54	61
5	34	42	50	56
6	33	40	47	53
7	32	39	45	50
8	31	37	43	48
9	31	36	41	46
10	30	35	40	45
11	30	35	39	44
12	29	34	38	43
13	29	33	38	42
14	29	33	37	41
15	28	32	36	40

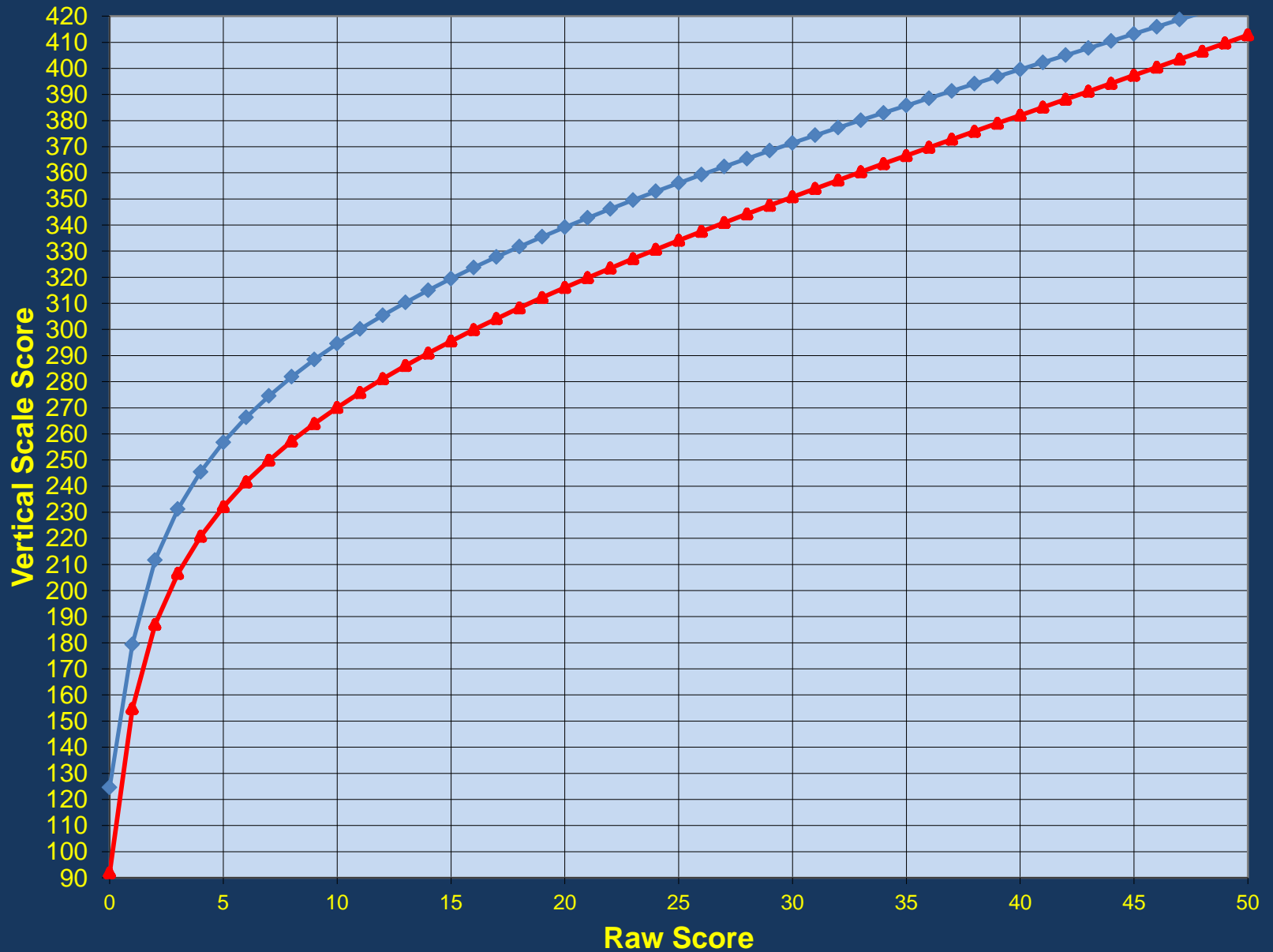
# Is this a Real Problem?

- This is a natural consequence of a nonlinear transformation
- This is not a problem inherent in the vertical scale
- Increases ( or decreases) at the extremes affect only a small percent of test takers
- However, it is an issue of FACE VALIDITY that should be addressed

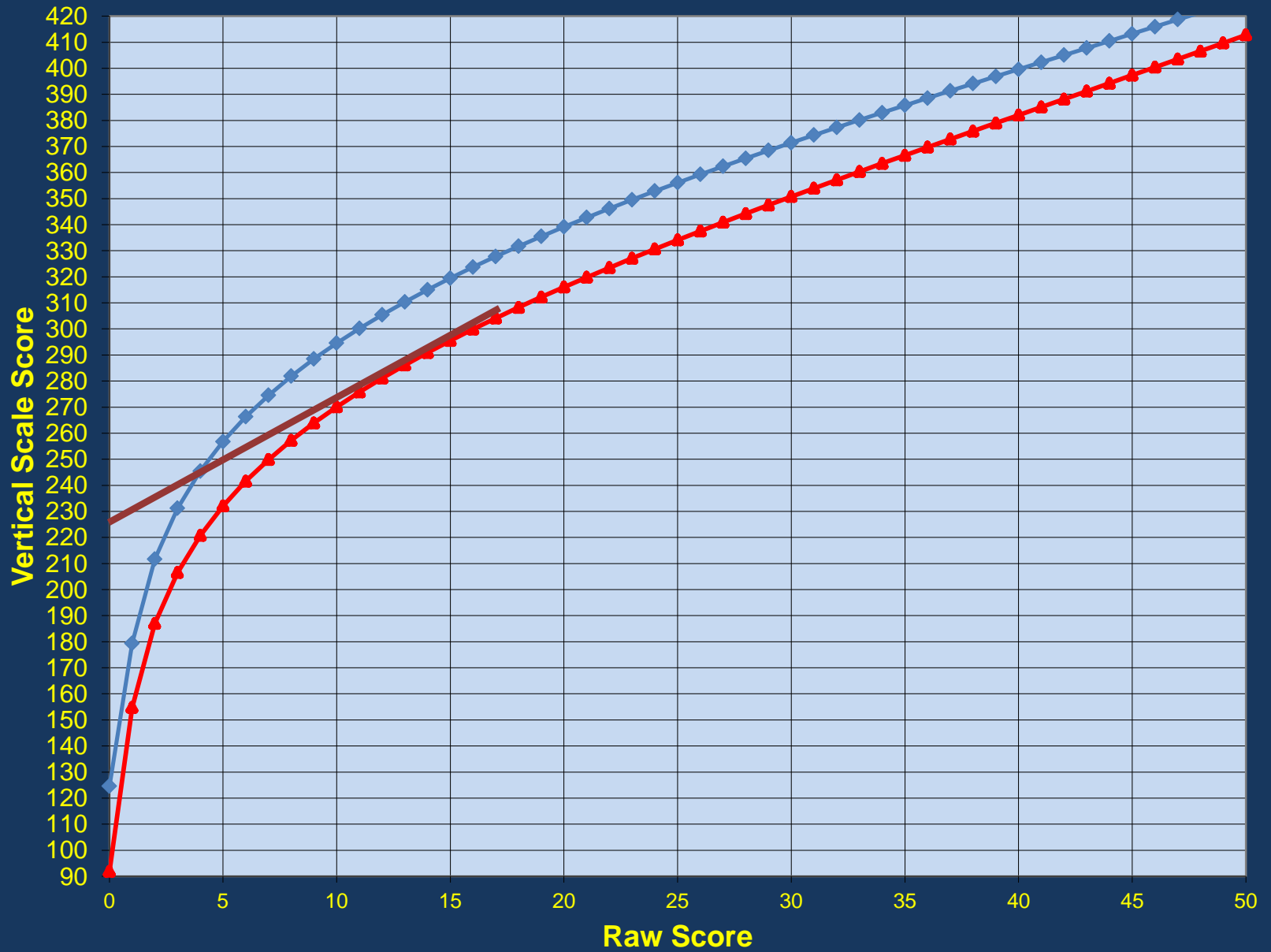
# Solution

- Adjustments are made to the two tails
- Linear interpolation at the extreme ends resolves the problem

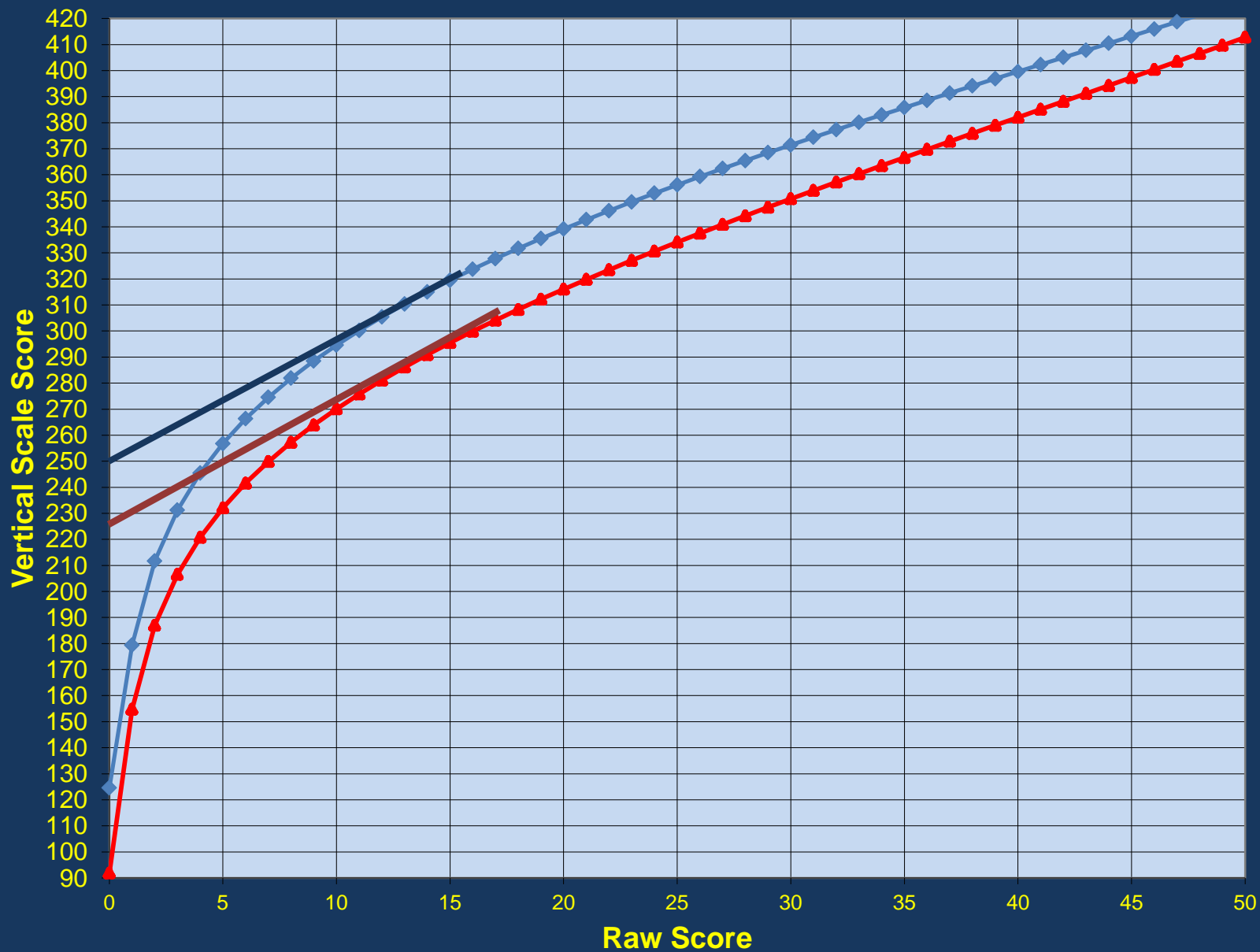
**Figure 3. ADJUSTMENT TO GRADES 3 AND 4 – (Lower End)**



**Figure 3. ADJUSTMENT TO GRADES 3 AND 4 – (Lower End)**



**Figure 3. ADJUSTMENT TO GRADES 3 AND 4 – (Lower End)**



# Steps

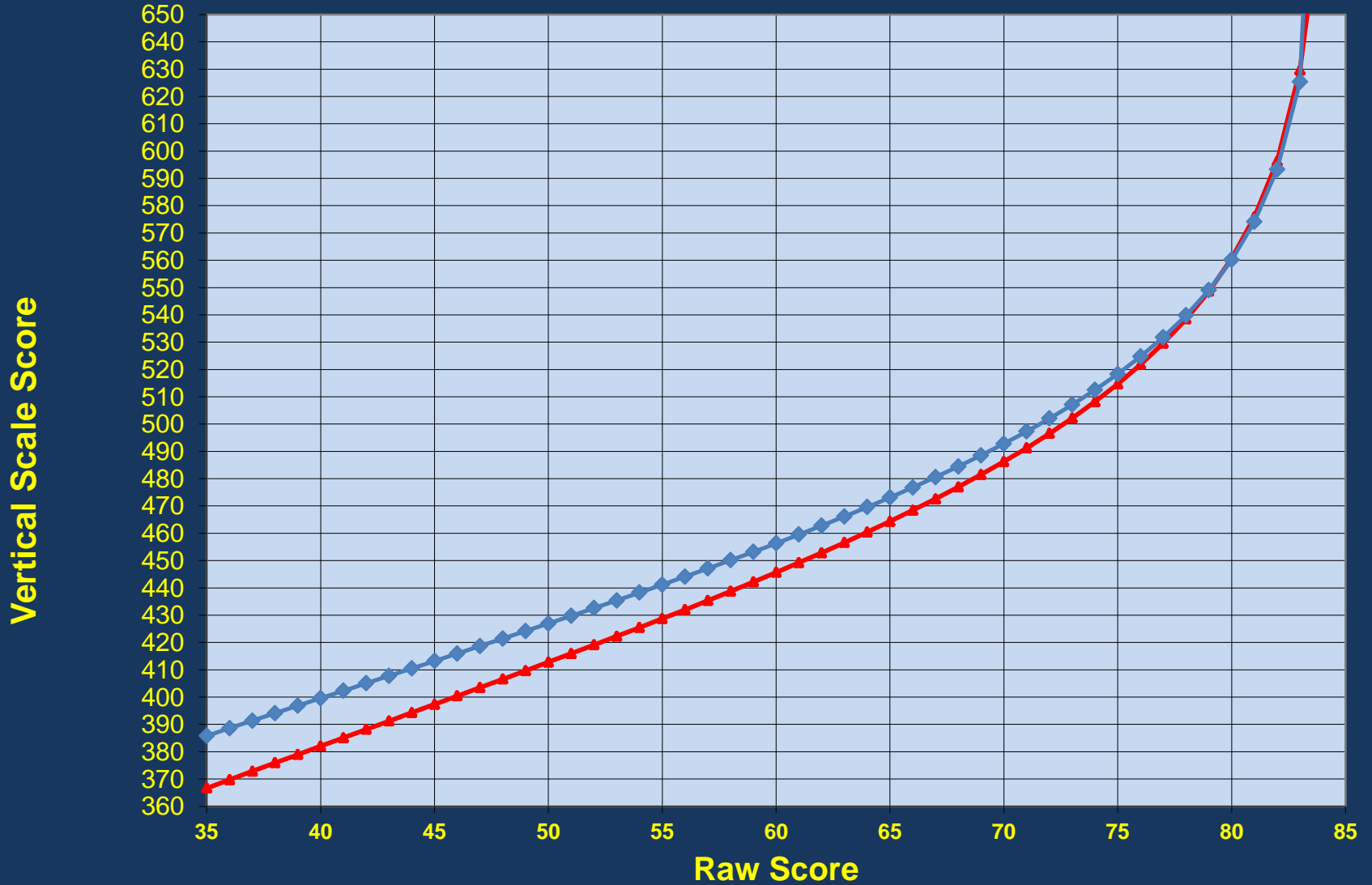
1. Choose a suitable region where the curve is almost linear – where the rate of change is almost a constant
2. Fit a regression line/ or an analytical line and extend this line.
3. Check to see if the growth (rate of change) is similar to that in the middle of the score range
4. Carry out the procedure across two grades, e.g., from Grade 3 to 4, checking to make sure the growth from one grade to the next is almost uniform across the raw score range
5. Keeping the Grade 4 line fixed, carry out the procedure across grades 4 and 5
6. If the change from grade 4 to 5 is not uniform, reexamine the growth from grades and modify the line(s) until the growths from grades 3 to 4 and 4 to 5 are “reasonable”.
7. Repeat this process across all the grades.



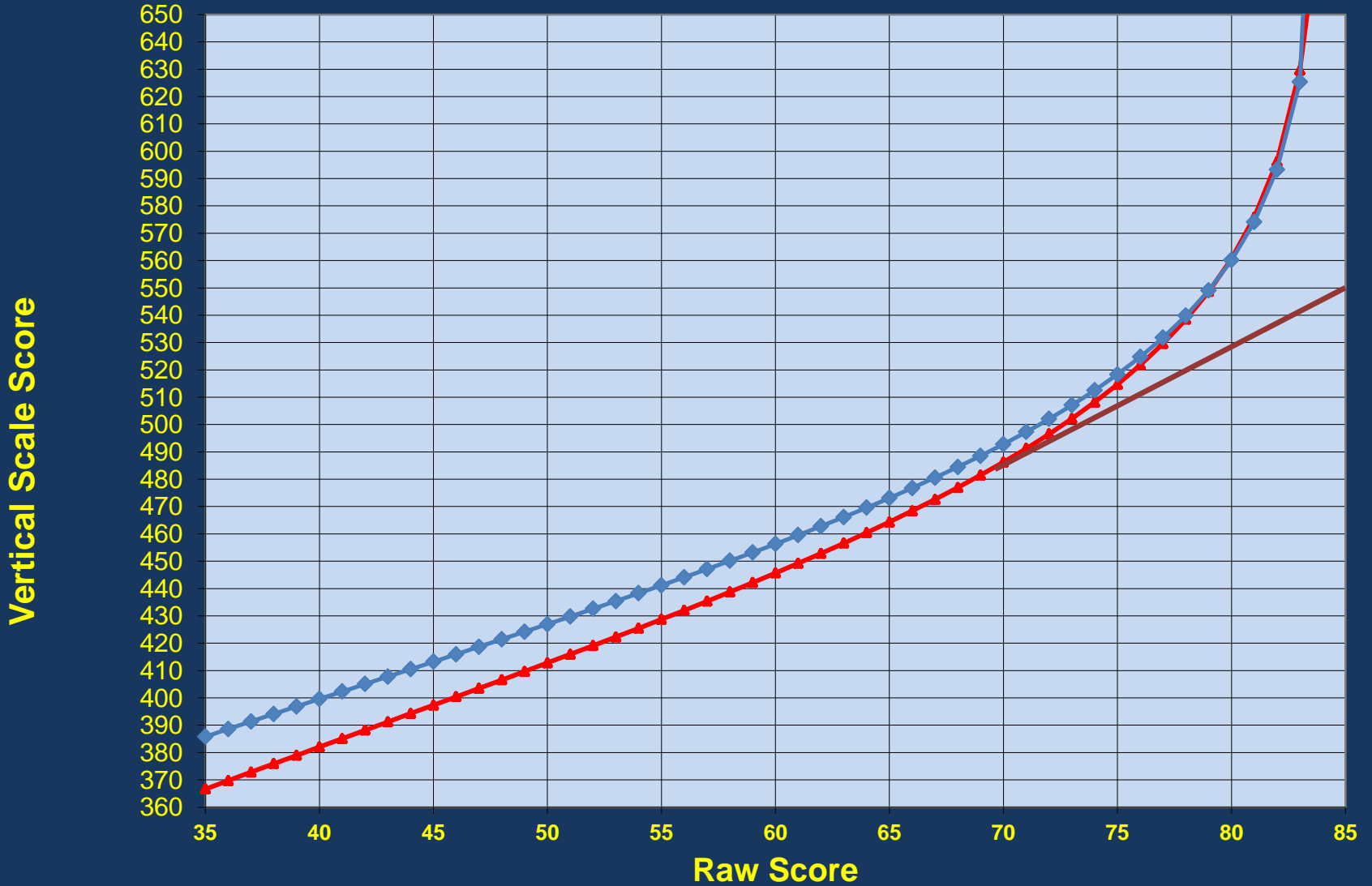
**Table 2. Gain in Adjusted Vertical Scale Score Corresponding to an Increase in Raw Score (LOWER END)**

Raw Score Grade 3	Gain using <i>Adjusted Vertical Scale Score</i> Grade 4			
	Raw Score +1	Raw Score +2	Raw Score +3	Raw Score +4
0	30	34	39	43
1	29	34	39	43
2	29	34	39	43
3	29	34	38	43
4	29	34	38	43
5	29	34	38	43
6	29	33	38	43
7	29	33	38	42
8	29	33	38	43
9	28	33	38	43
10	28	33	38	43
11	28	33	38	42
12	29	33	38	42
13	28	33	37	41
14	28	32	37	40
15	28	32	36	40

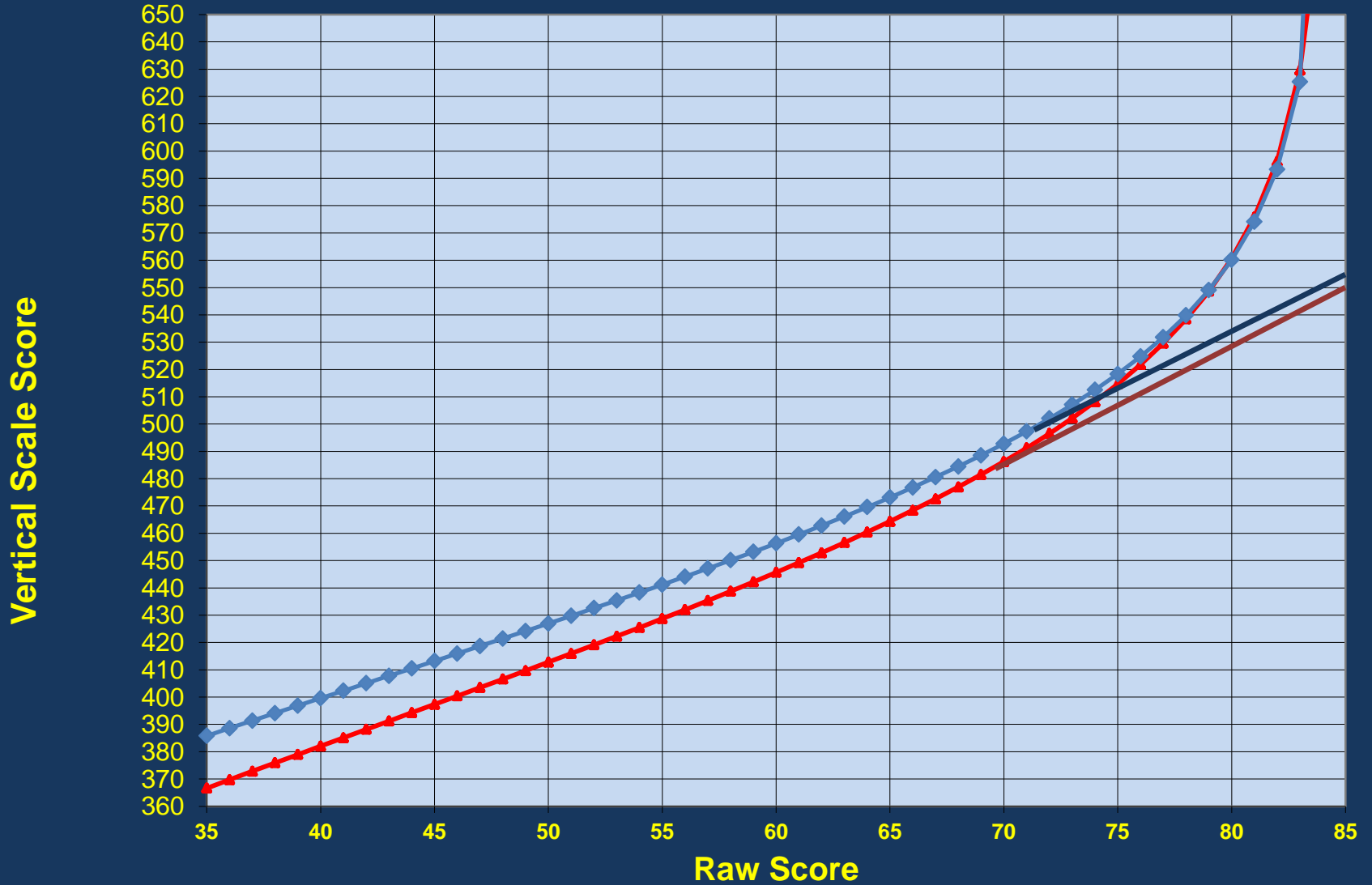
**Figure 4. Adjustments to Grades 3 and 4 (Upper End)**



**Figure 4. Adjustments to Grades 3 and 4 (Upper End)**



**Figure 4. Adjustments to Grades 3 and 4 (Upper End)**



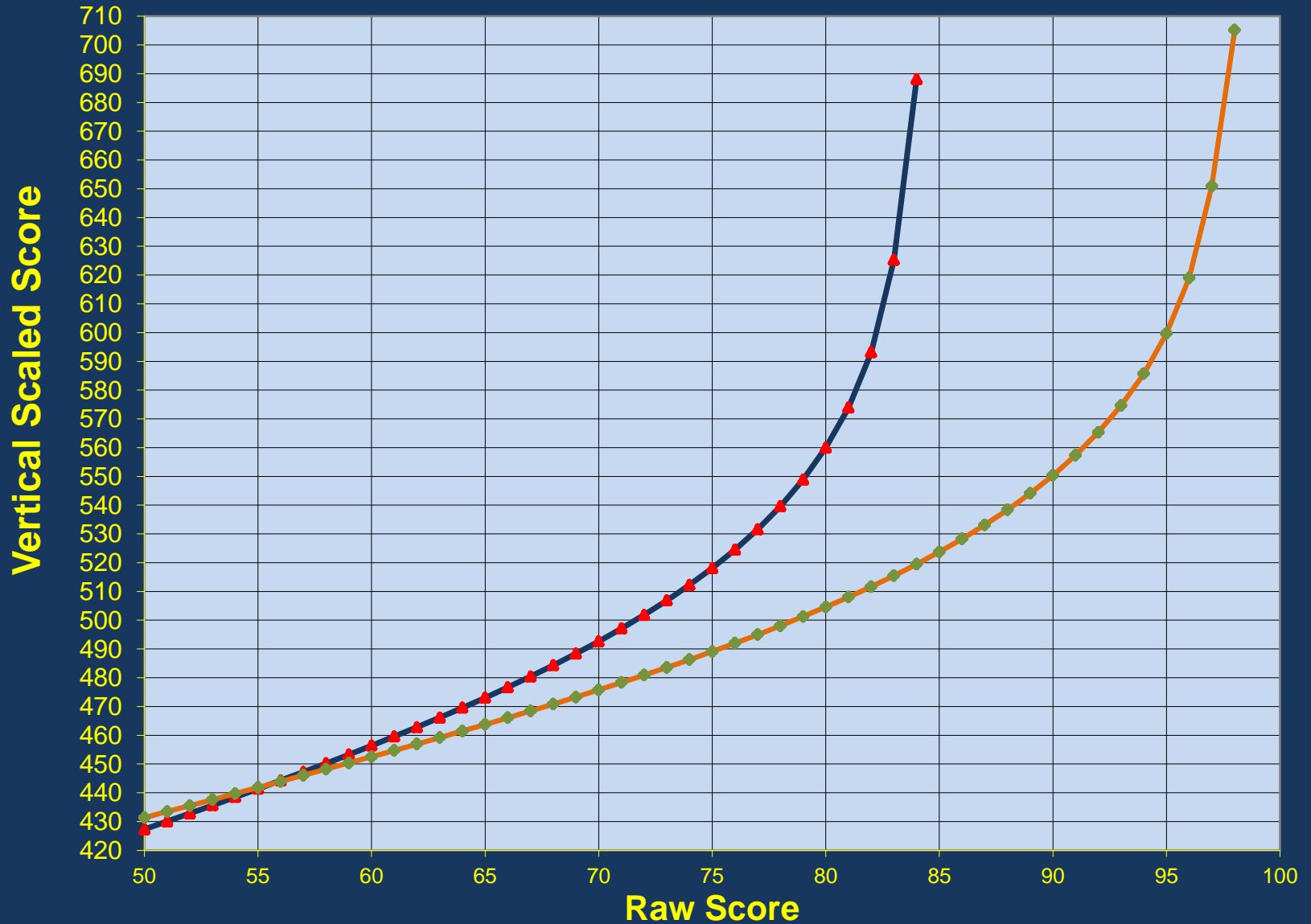
**Table 4. Gain in Adjusted Vertical Scale Score Corresponding to an Increase in Raw Score (Upper End)**

Raw Score Grade 3	Gain using Adjusted Vertical Scale Score at the upper end of the scale			
	Raw Score +1	Raw Score +2	Raw Score +3	Raw Score +4
70	11	15	20	24
71	11	15	20	24
72	11	16	20	24
73	11	16	20	24
74	11	16	20	25
75	11	16	20	25
76	11	16	20	25
77	11	16	20	25
78	12	16	20	25
79	12	16	20	25
80	12	16	21	25
81	12	16	21	
82	12	16		
83	12			
84*				

# A Difficulty

- A difficulty arises when test in one grade has a different number of items than the test in another grade
- What happens when Grade 4 test has 84 items but Grade 5 test has 98 items?
- The curves will cross

**Figure 4. Relationship between Vertical Scaled Score and Raw Score: Grade 4 to Grade 5 (Upper End - Reading)**



- The curves cross at the raw score value of 56.
- A student with a raw score of 70 in Grade 4 and 70 in Grade 5 will show a decline in the VSS of 15 VSS points ( $490 - 475 = -15$ ).
- A score of 70 in Grade 4 corresponds to percent correct score of **83.3**
- A score of 70 in Grade 5 corresponds to percent correct score of **71.4**

A Difficulty? – Not really!

Report Percent Correct Scores



# CONCLUSIONS

- IRT scaling results in a nonlinear relationship between raw score and Vertically Scaled Score
- Small changes in raw scores at the extreme result in large changes in VSS at the extremes
- This is not a theoretical drawback of the VSS, but a perception problem for the test users
- This is an issue of face validity
- Problem is addressed by simple adjustments at the extremes

# Conclusions

- We do not recommend describing growth with reference to raw scores
- If observed scores on the test have to be reported, use Percent Correct Scores. This avoids appearances of problems when tests have different number of raw score points

# A Posteriori Possible Literary Titles

- TALE OF TWO TAILS ?
- THE TEMPEST in a tea pot ?
- MUCH ADO ABOUT NOTHING ?