Patient reported Outcomes: Misinference from ordinal scales?

Mike Horton
Research Assistant, PhD Student

Alan Tennant PhD
Professor of Rehabilitation Studies

Psychometric Laboratory for Health Sciences
Department of Rehabilitation Medicine,
University of Leeds, UK

MRC Clinical Trials Methodology Conference
October 4-5, 2011; Bristol, UK.

UNIVERSITY OF LEEDS

Patient Report Outcome Measures (PROM)

- PROs/PROMs
- Often used as outcomes in clinical trial/medical outcome studies
- Summated set of items provide ORDINAL magnitude of the respondents level of the trait.

Outcome studies

Many rely on the calculation of:

- Means + SDs
- Change Scores
- Minimally Important Difference
    - MCID/MCII/MID
- Effect Sizes

Misinference from Ordinal Scales

- Ordinal scales of measurement do not support the mathematical operations needed to calculate means and standard deviations

Merbitz C, Morris J, Grip JC. Arch Phys Med Rehabil. 1989: 70:308-312

Ordinal scale and statistics in medical research.

- "When several items are measured on ordinal scales it is far from certain that the sum of scores has even ordinal properties"

Forrest M, Andersen B. Ordinal scale and statistics in medical research. Br Med J (Clin Res Ed). 1986: 292(6519): 537-8

Ordinal scale and statistics in medical research.

- A survey of the 1982 editions of 12 medical journals showed that in at least 70% of 175 papers employing ordinal measurement scales, statistical methods were used, which do, in fact, assume a more refined measurement scale.

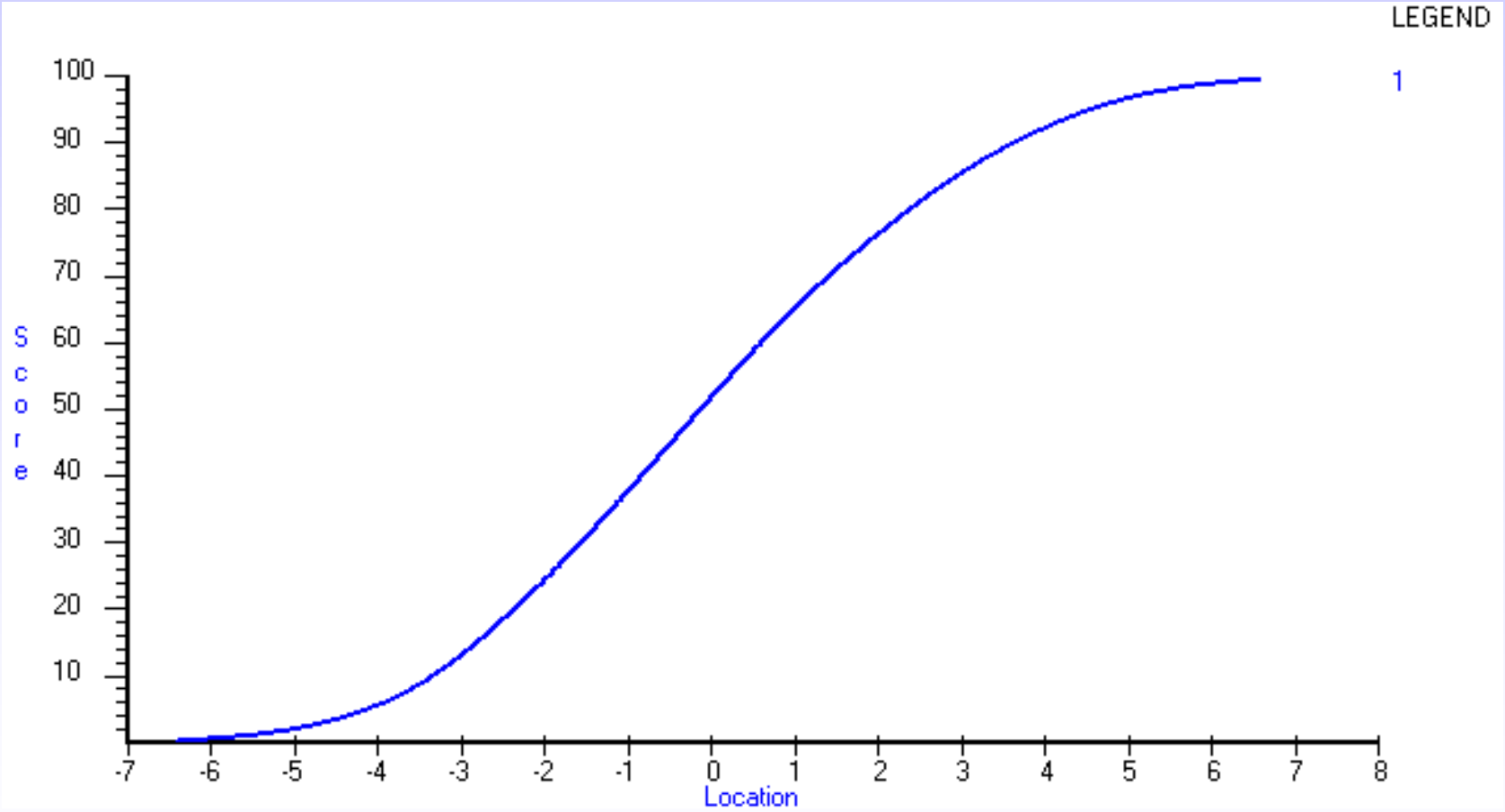
Forrest M, Andersen B. Ordinal scale and statistics in medical research. Br Med J (Clin Res Ed). 1986: 292(6519): 537-8

Scale Warning

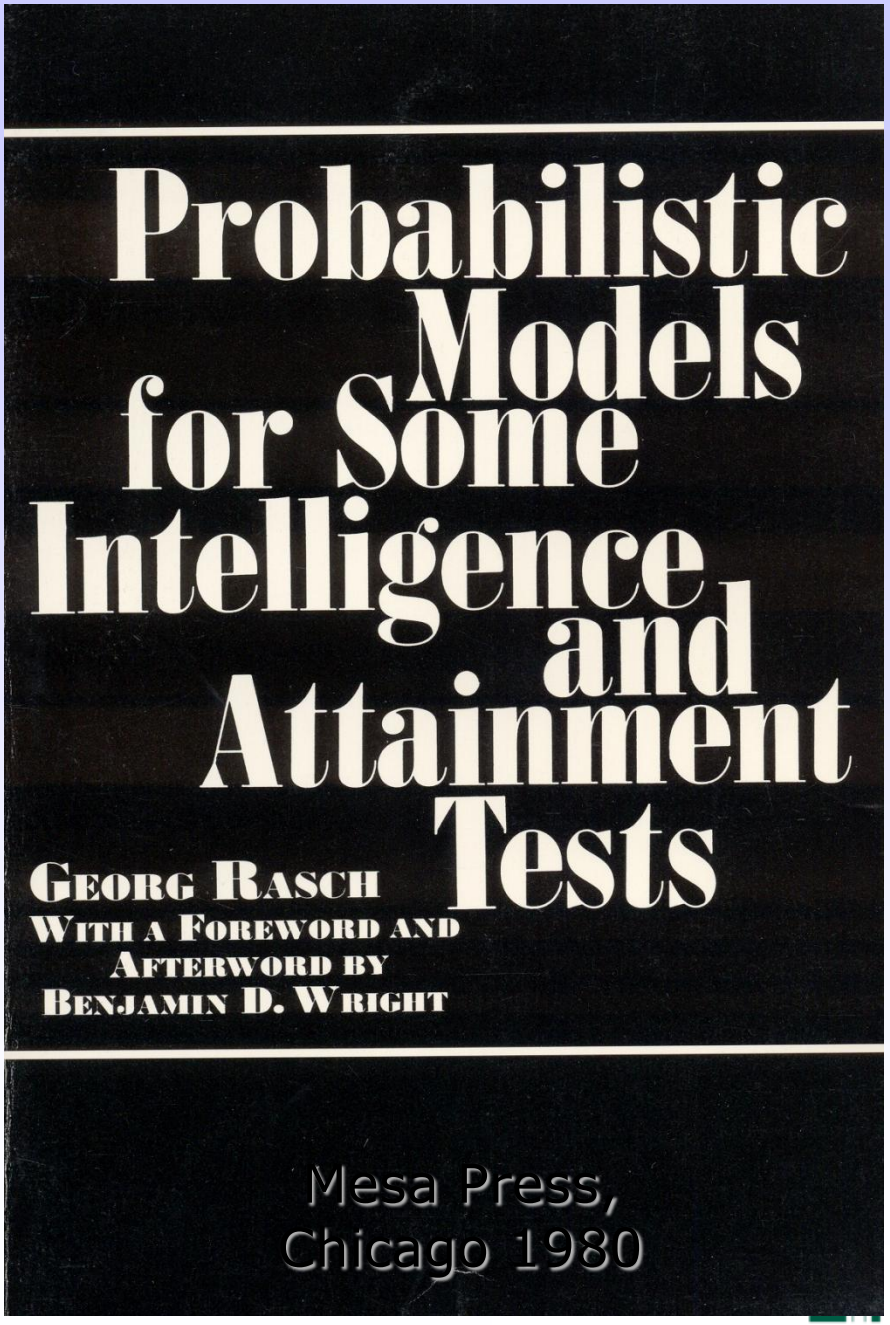
Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts

Stuki G, Daltroy L, Katz JN, Johannesson M, Liang MH. Journal of Clinical Epidemiology 1996; 49:711-717

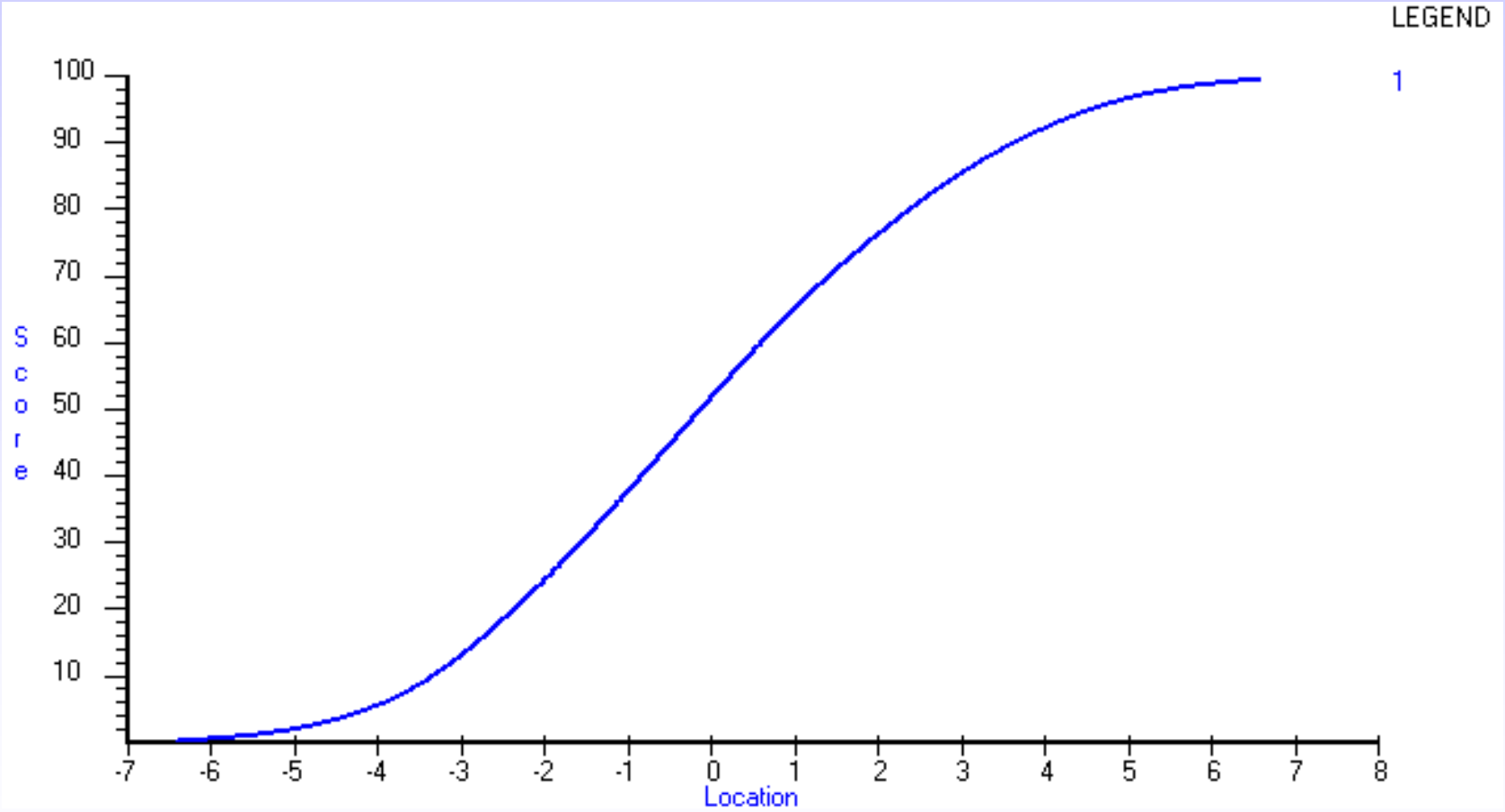
Transformation from Ordinal to Interval



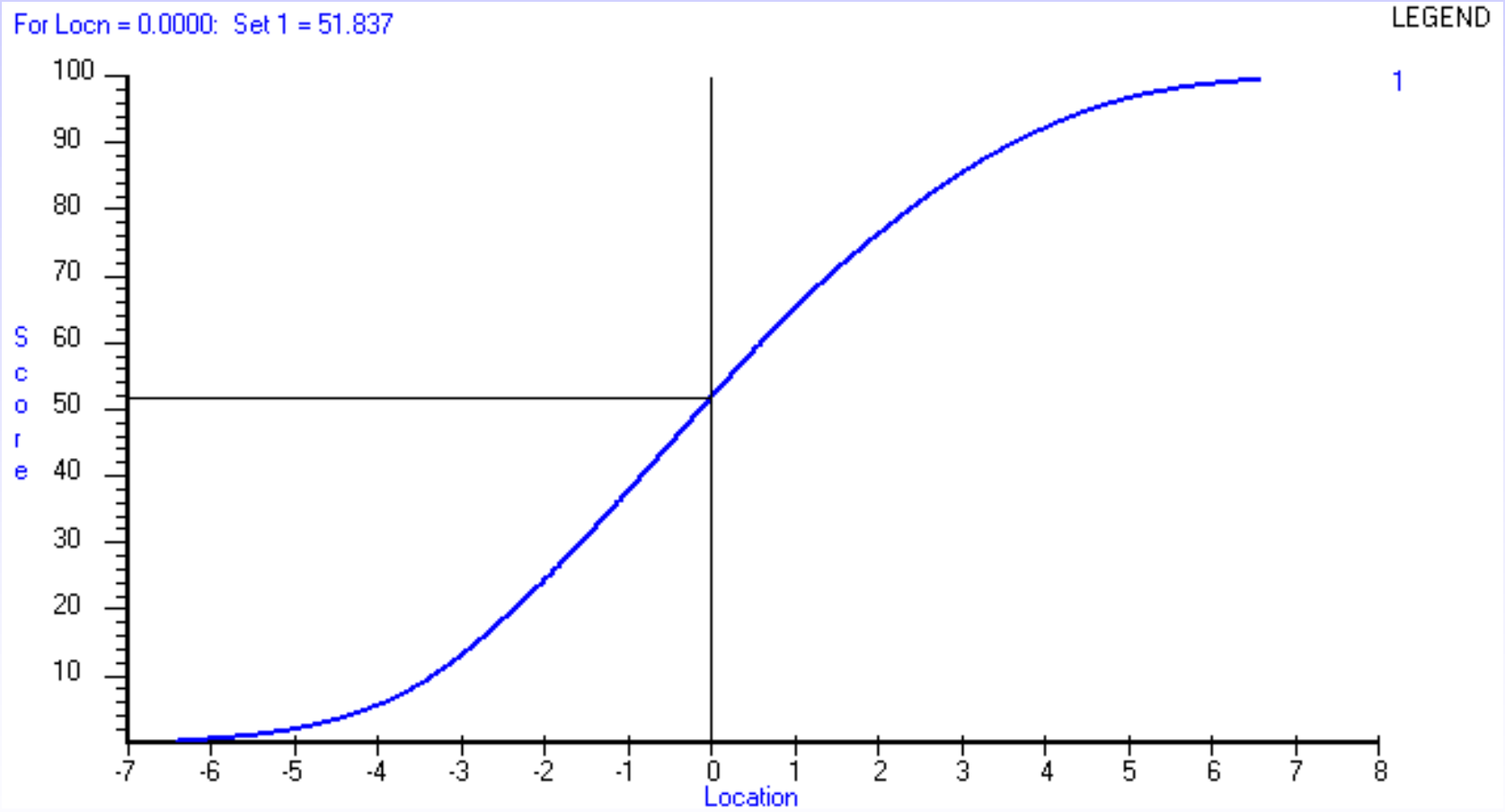
Georg Rasch (1901-1980)



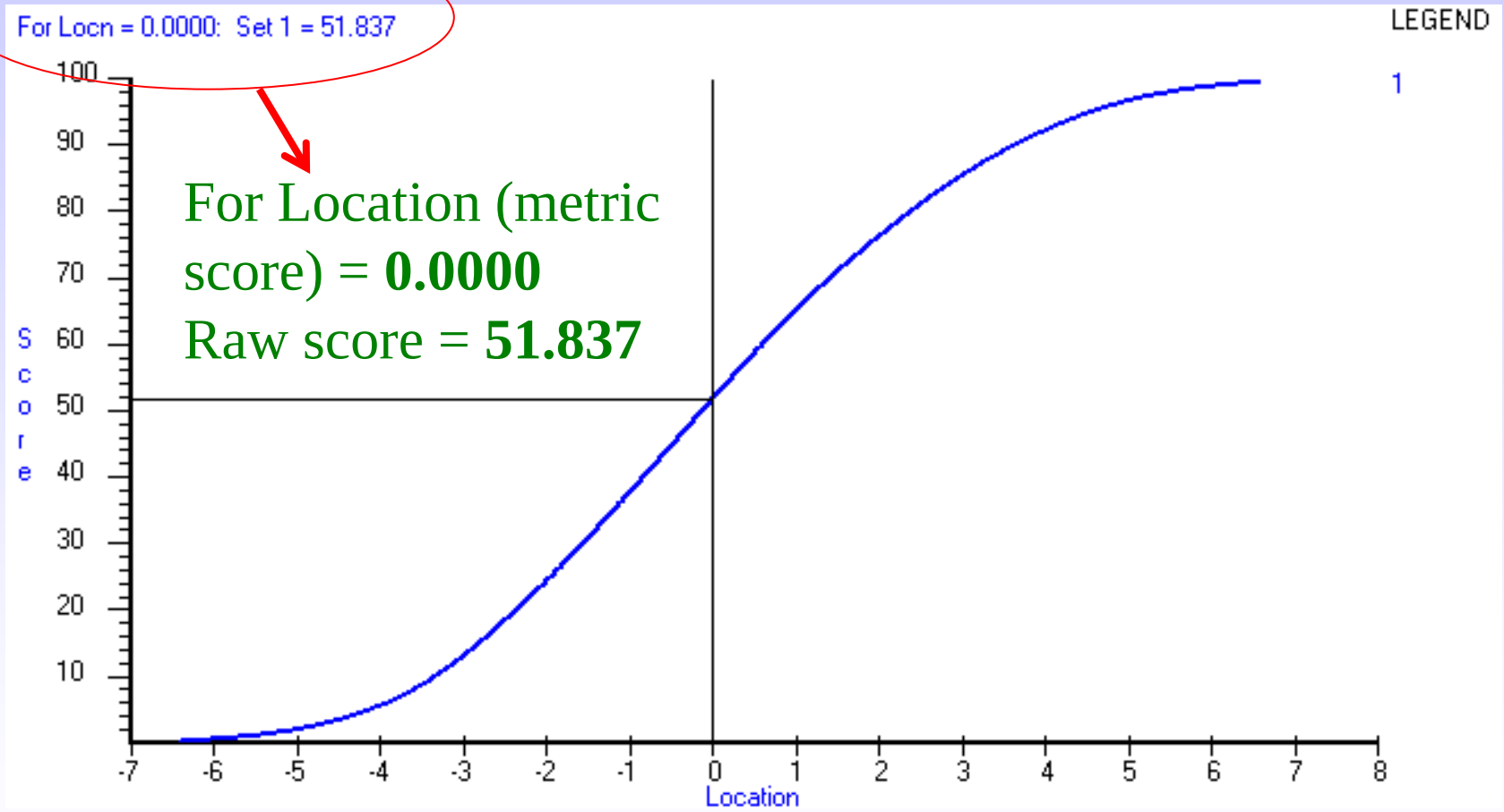
Example 1: Plateau?



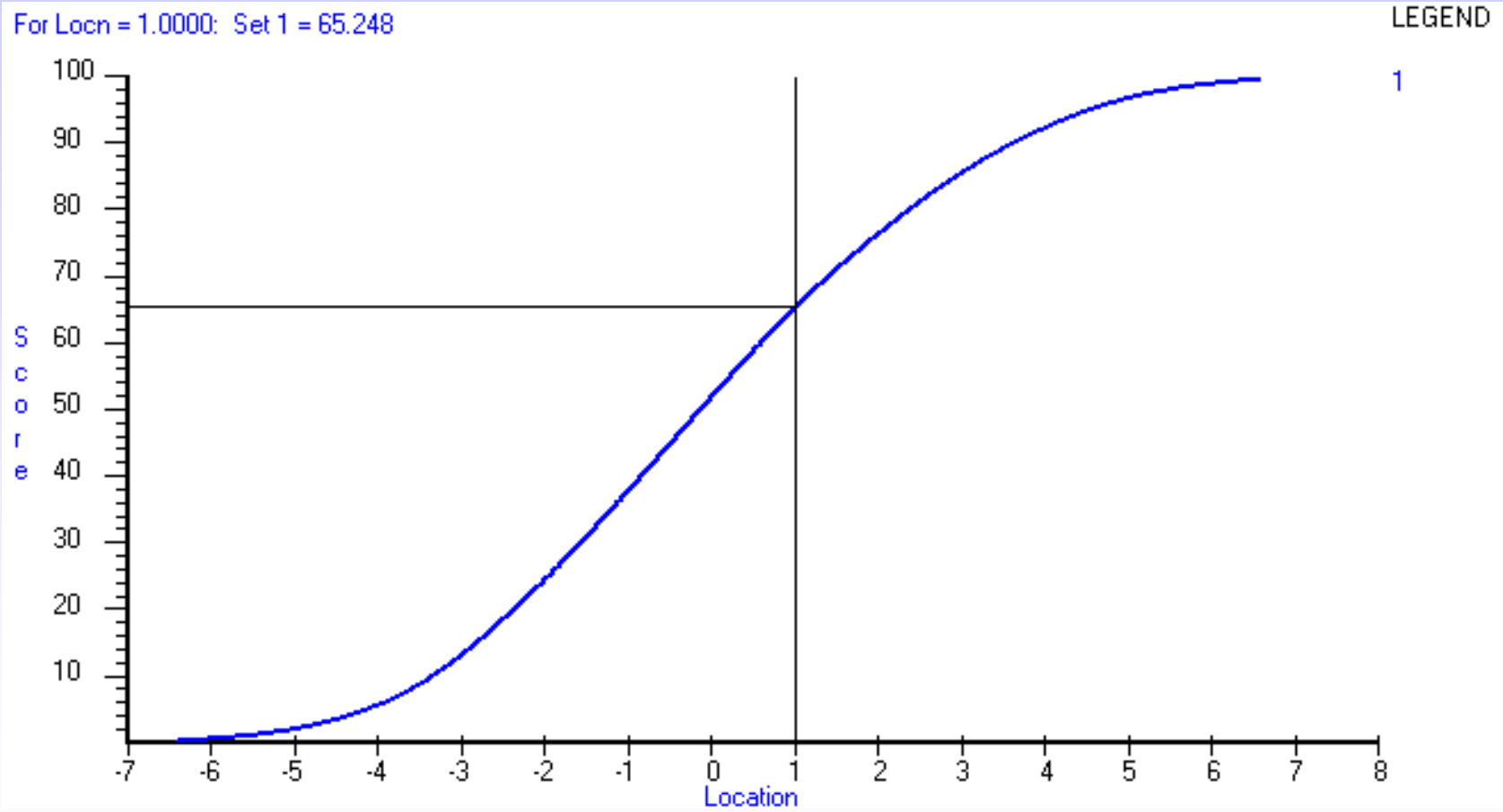
Plateau?


For Locn = 0.0000: Set 1 = 51.837
LEGEND
1
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

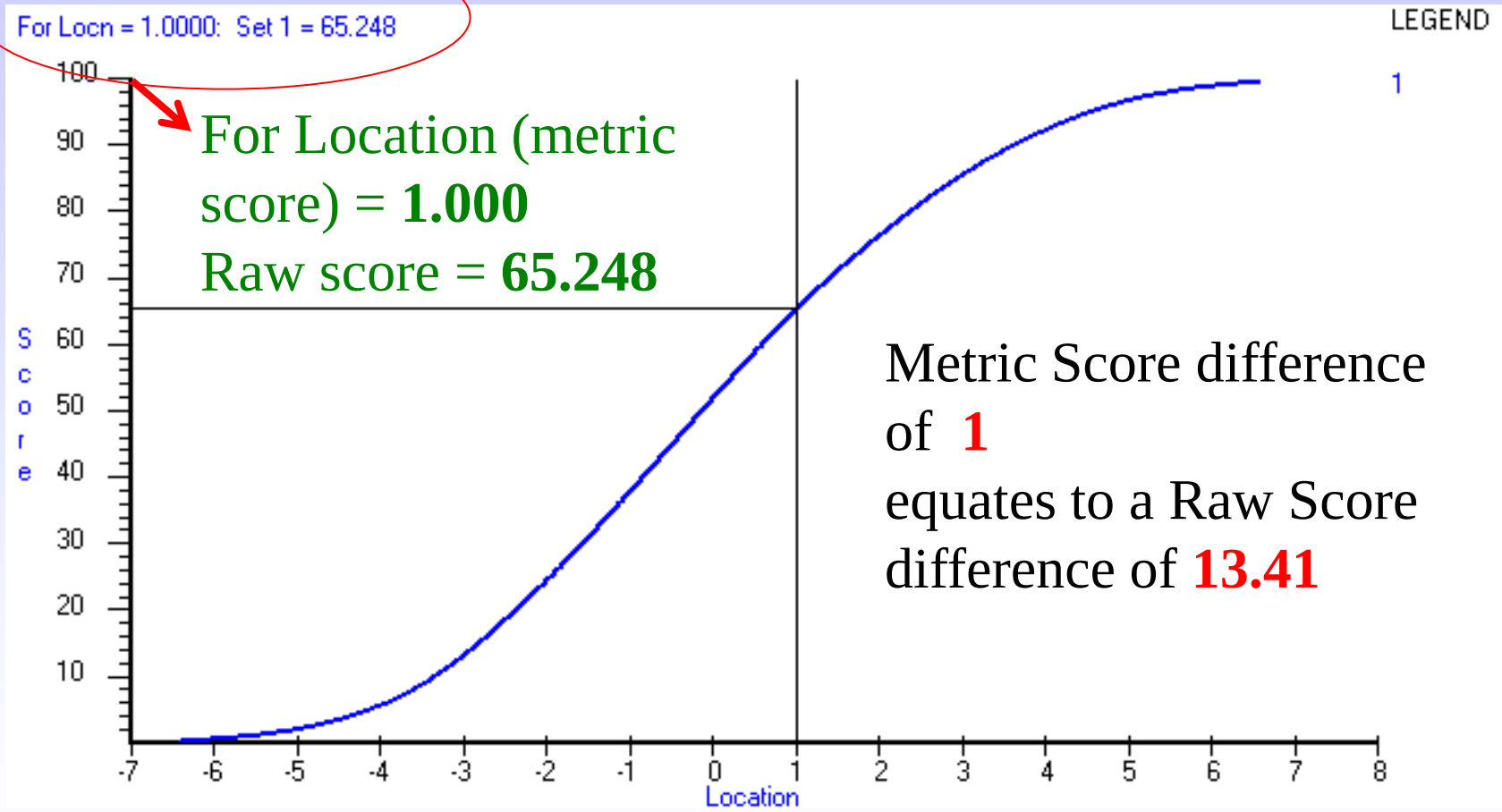
Plateau?


For Locn = 0.0000: Set 1 = 51.837
LEGEND
1
For Location (metric score) = **0.0000**
Raw score = **51.837**
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

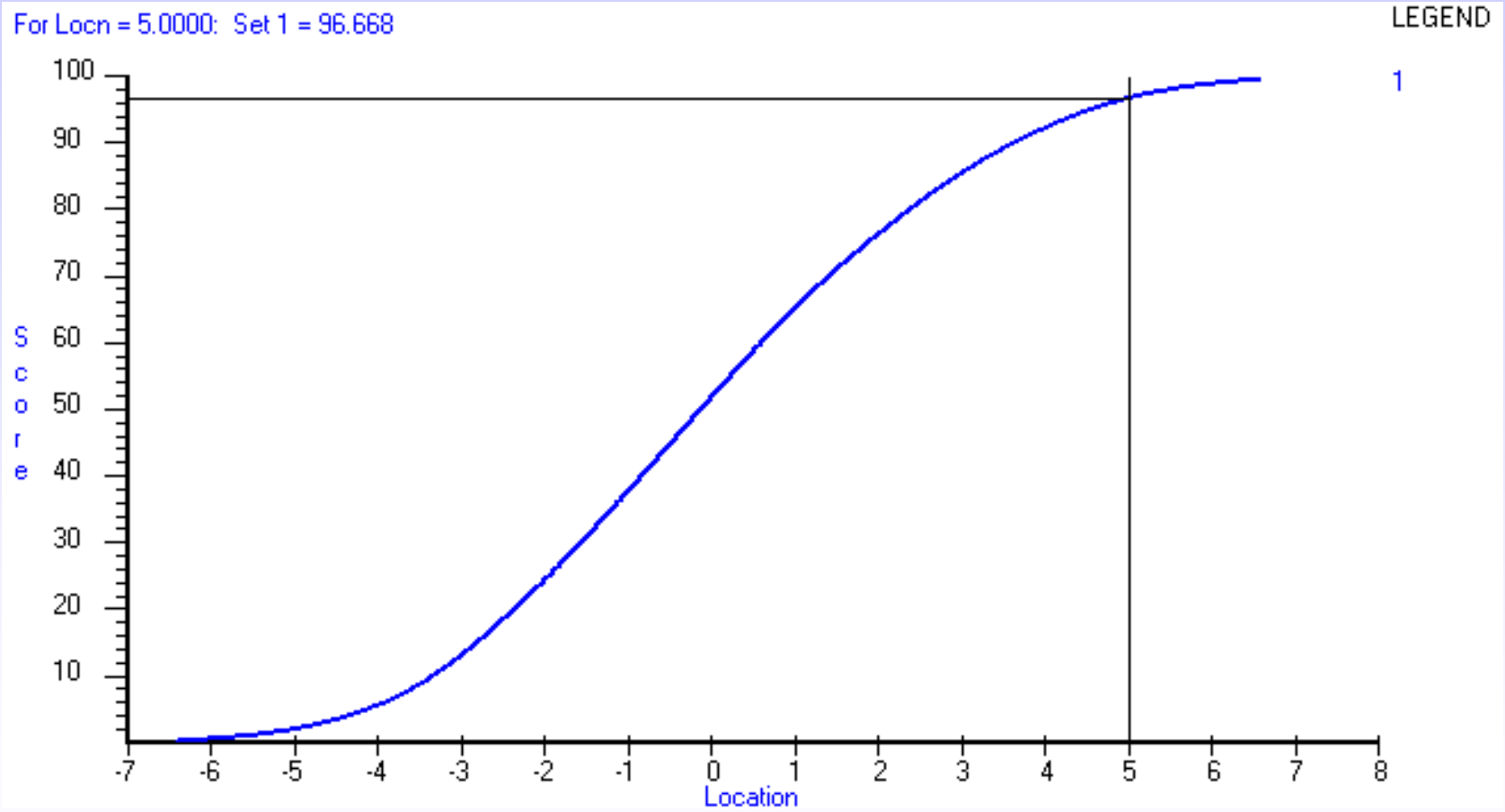
Plateau?


For Locn = 1.0000: Set 1 = 65.248
LEGEND
1
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

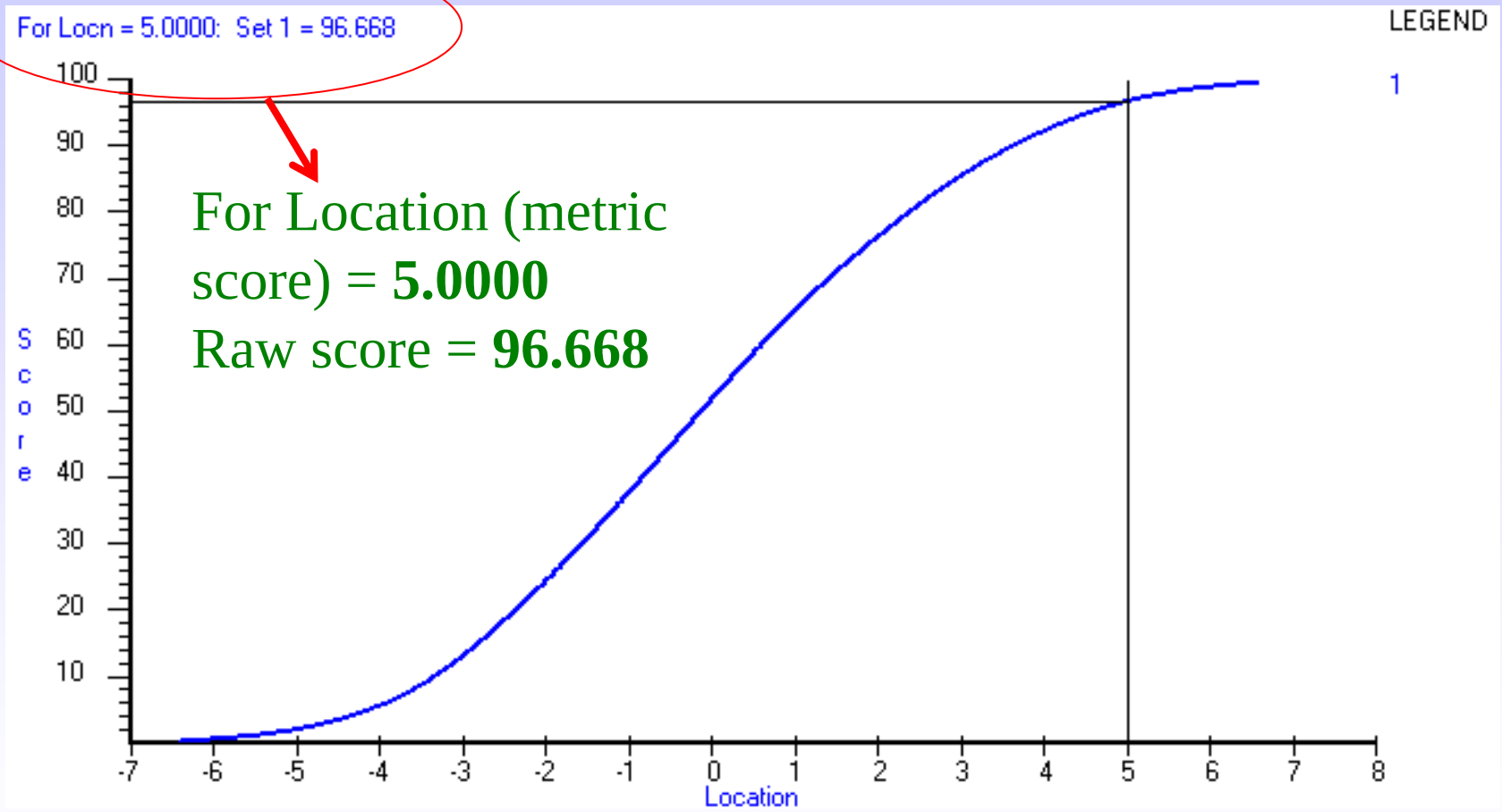
Plateau?


For Locn = 1.0000: Set 1 = 65.248
LEGEND
1
For Location (metric score) = **1.000**
Raw score = **65.248**
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location
Metric Score difference of **1**
equates to a Raw Score difference of **13.41**

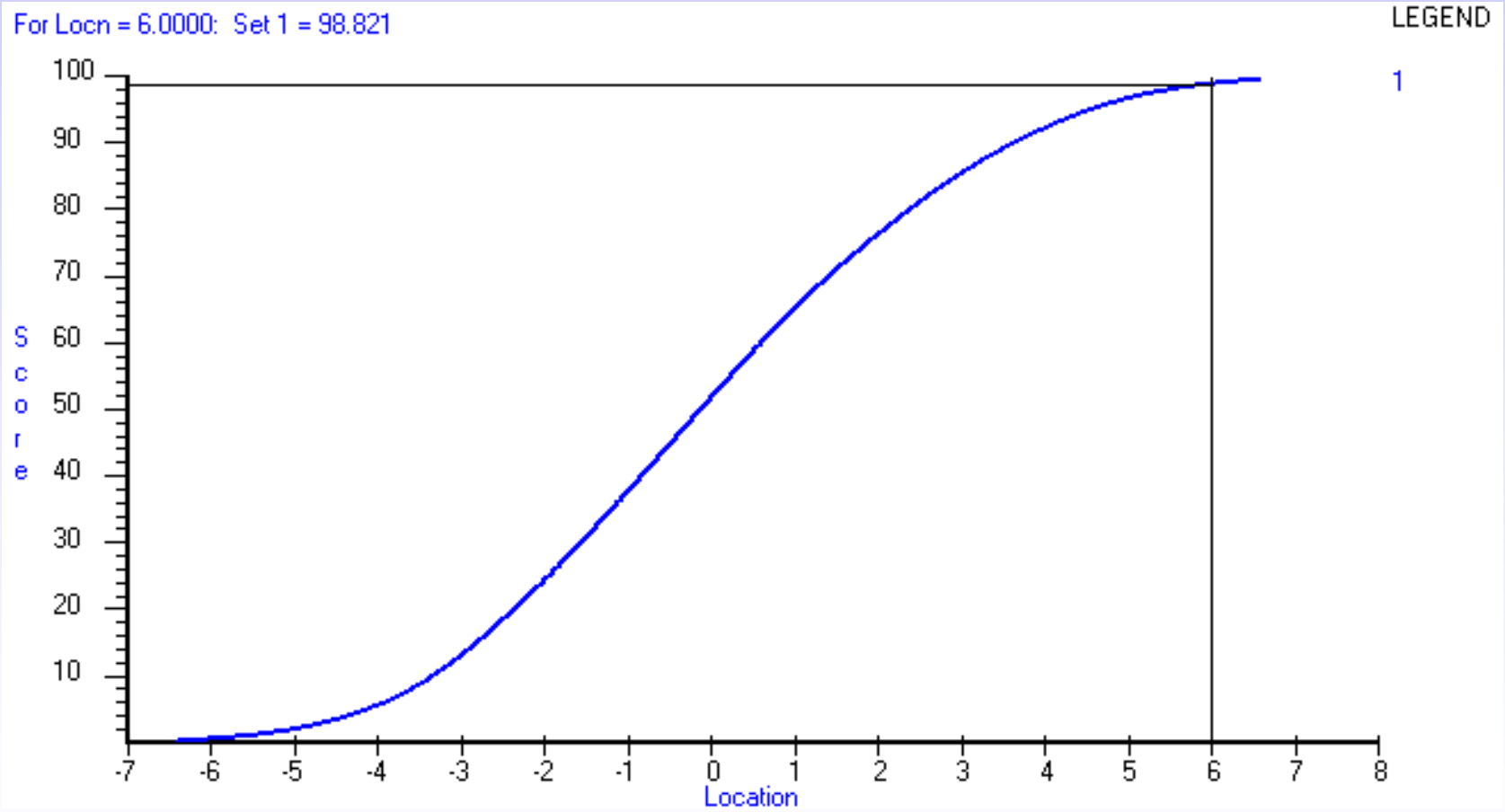
Plateau?


For Locn = 5.0000: Set 1 = 96.668
LEGEND
1
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

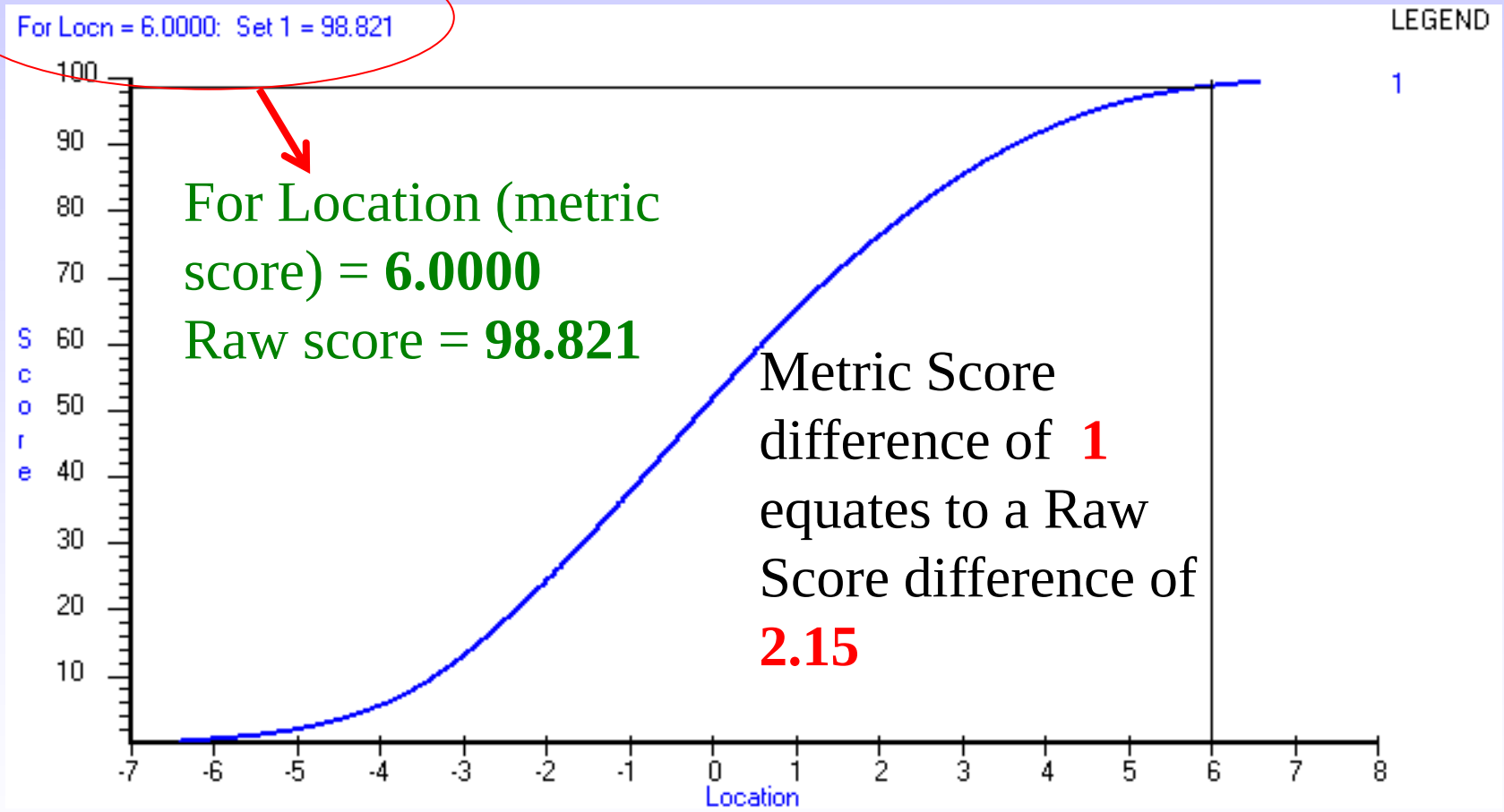
Plateau?


For Locn = 5.0000: Set 1 = 96.668
LEGEND
1
For Location (metric score) = **5.0000**
Raw score = **96.668**
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

Plateau?

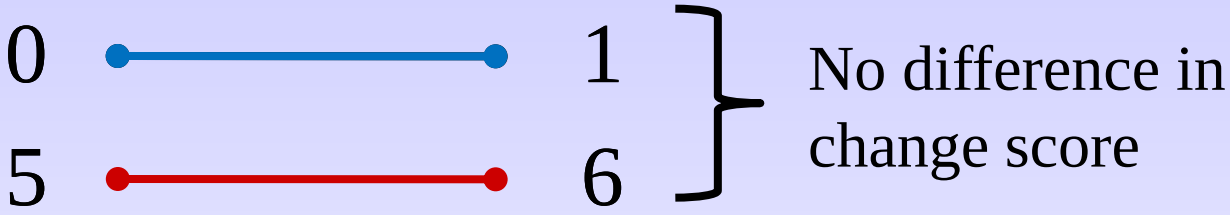

For Locn = 6.0000: Set 1 = 98.821
LEGEND
1
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

Plateau?

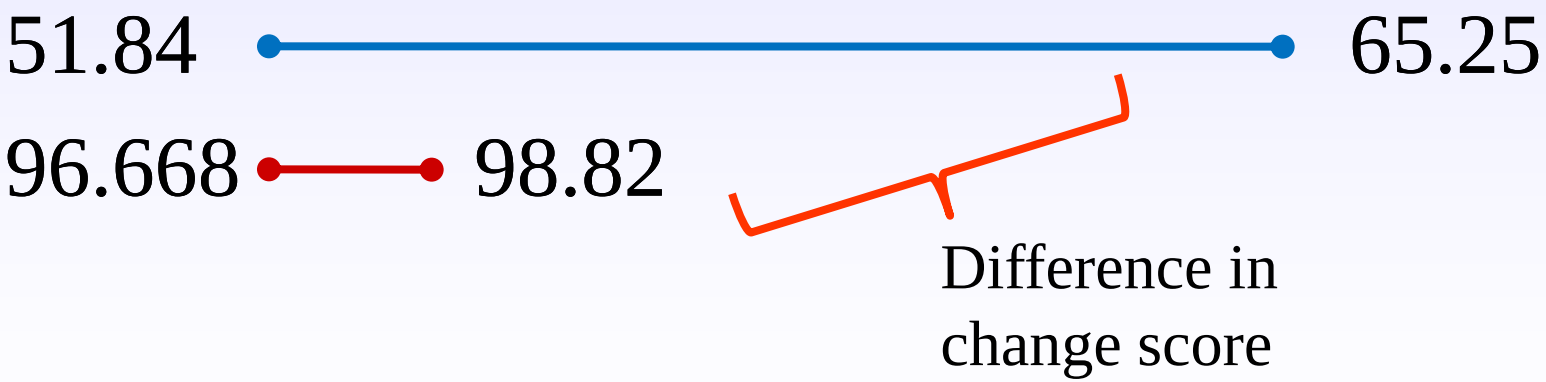

For Locn = 6.0000: Set 1 = 98.821
LEGEND
1
For Location (metric score) = 6.0000
Raw score = 98.821
Metric Score difference of 1 equates to a Raw Score difference of 2.15
Score
Location
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8

Difference?

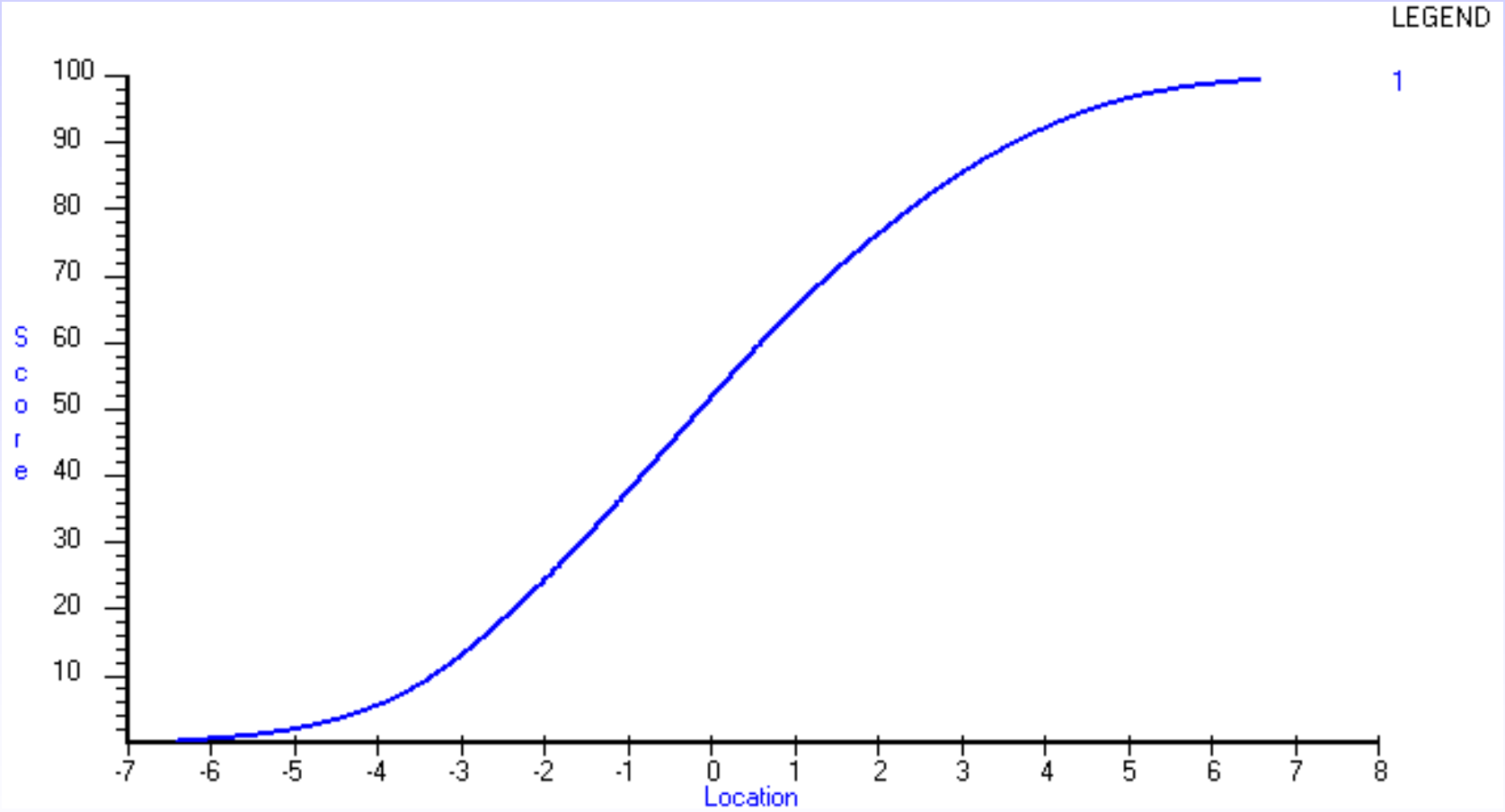
With Metric Scale Scoring:



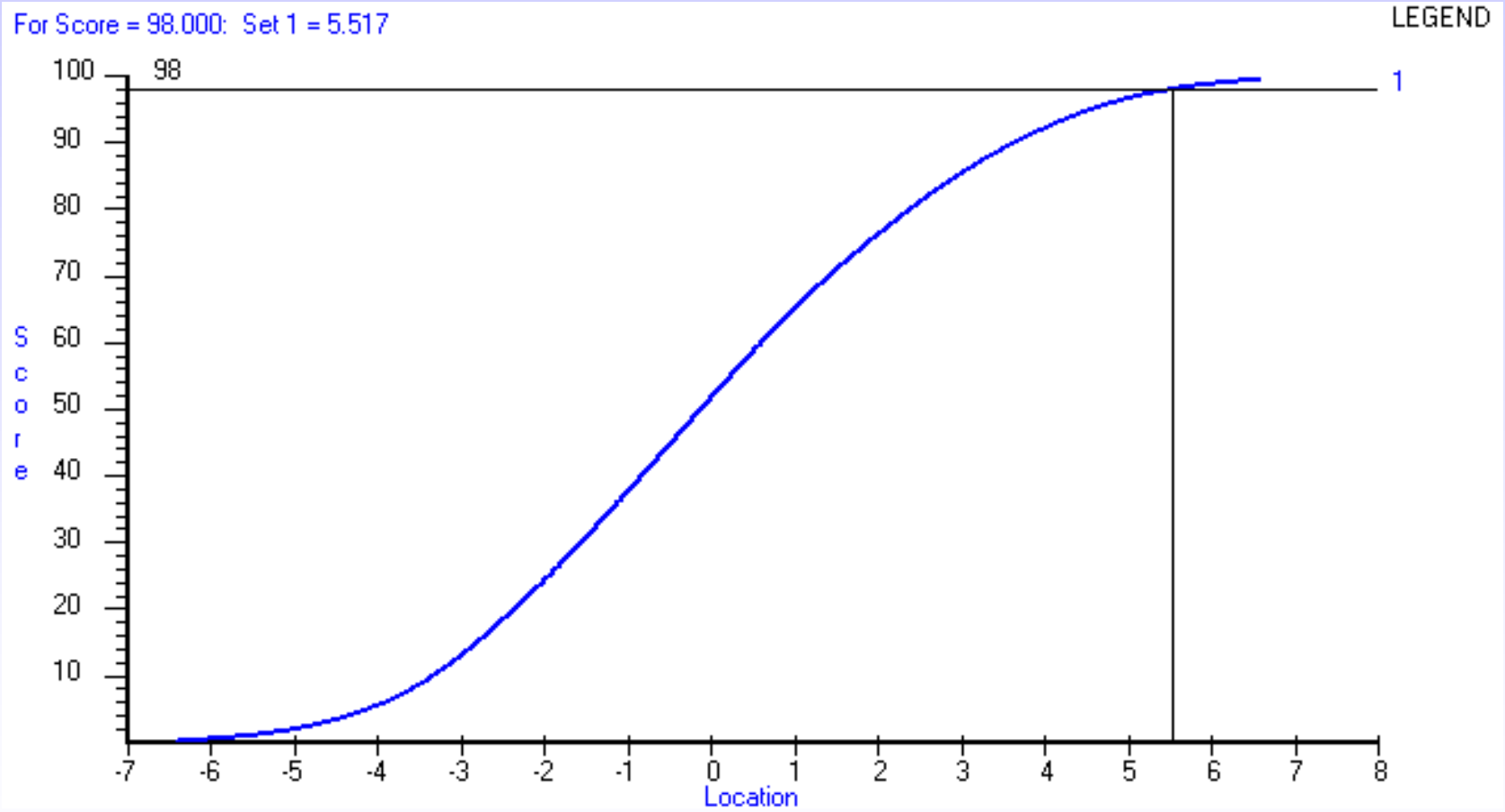
With Raw Ordinal Scale Scoring



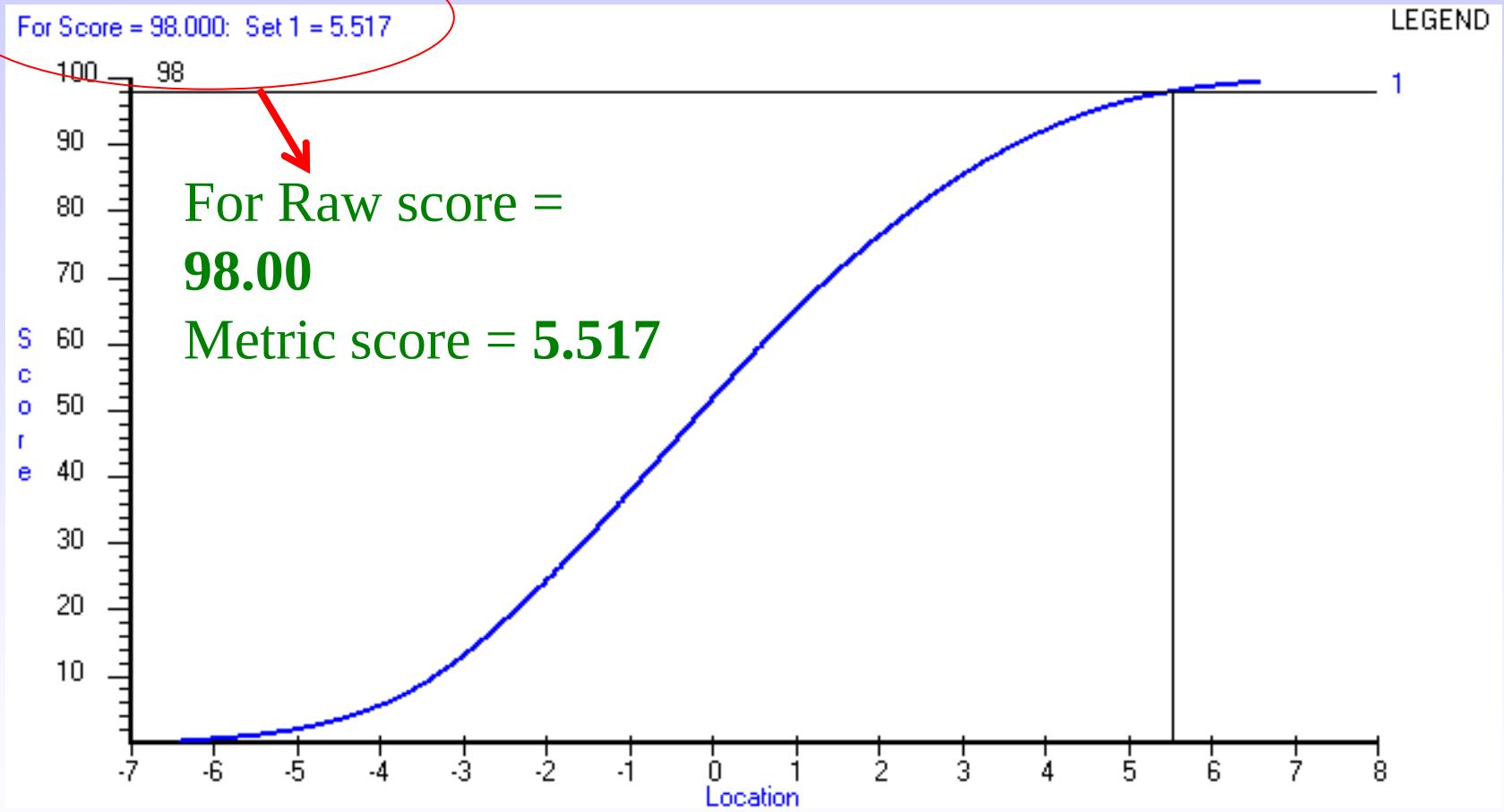
Example 2: MID of - 5?



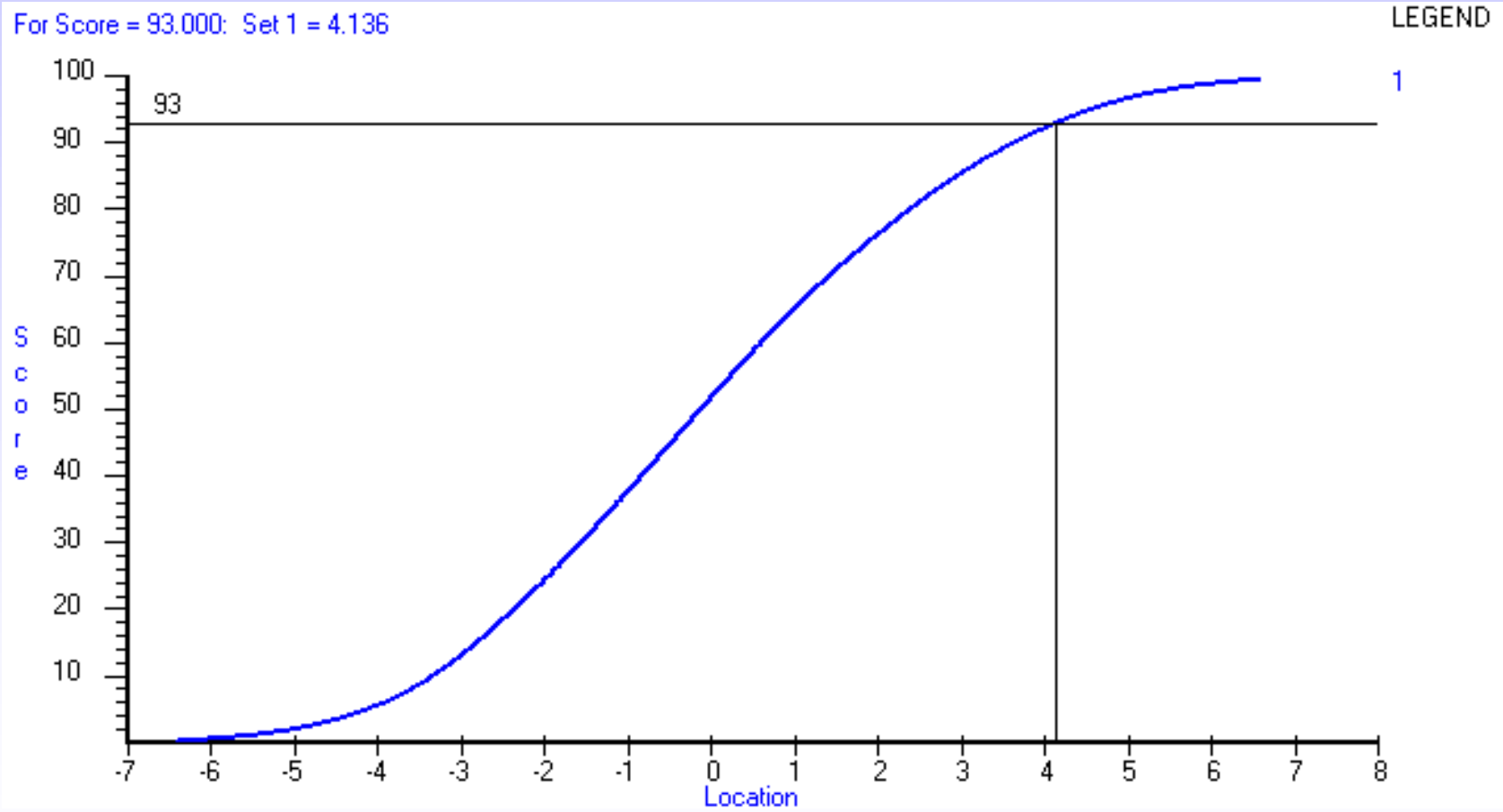
MID of - 5?


For Score = 98.000: Set 1 = 5.517
LEGEND
98
1
100
90
80
70
60
50
40
30
20
10
Score
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

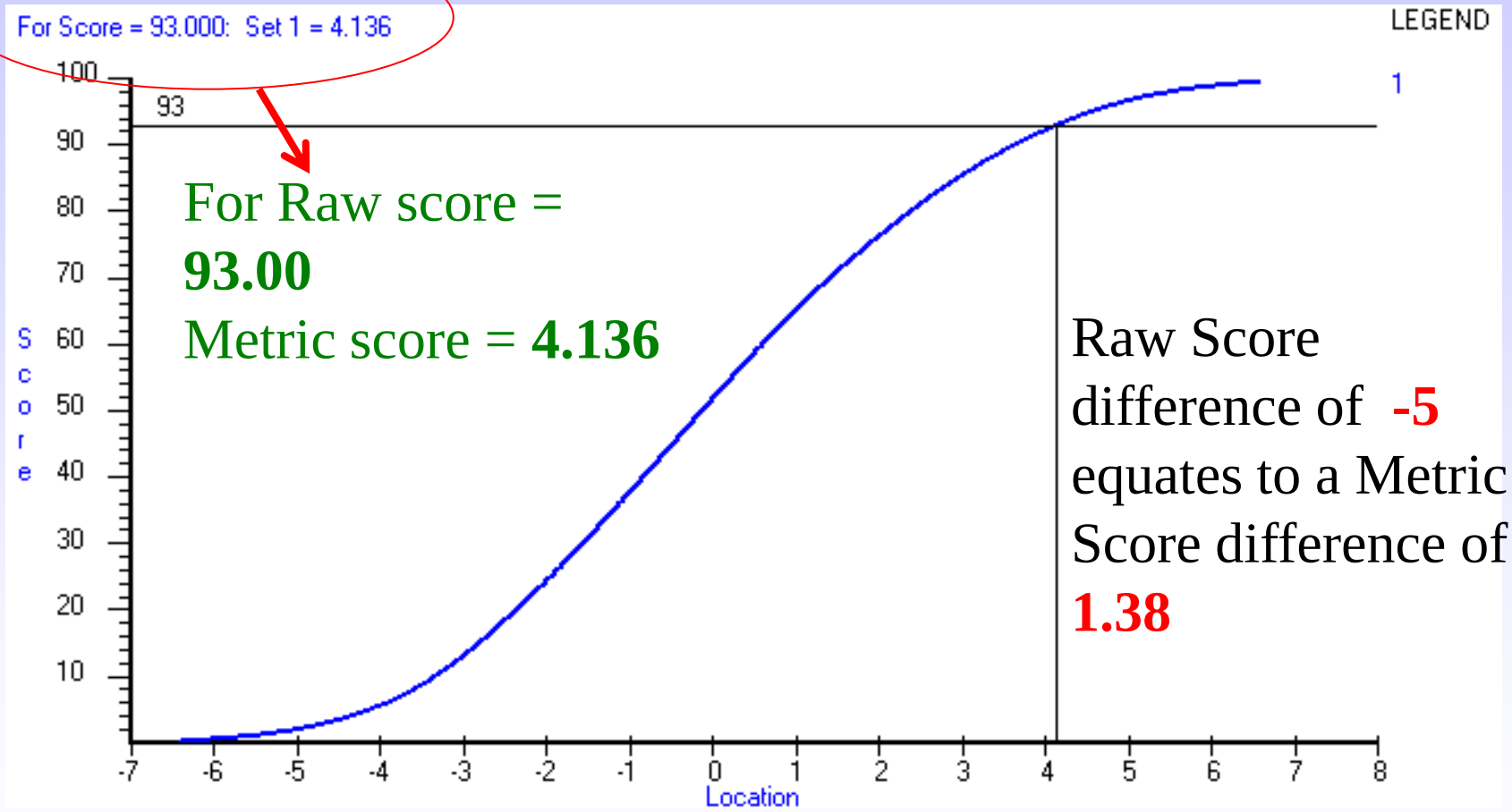
MID of - 5?


For Score = 98.000: Set 1 = 5.517
LEGEND
1
98
100
90
80
70
60
50
40
30
20
10
Score
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location
For Raw score =
98.00
Metric score = 5.517

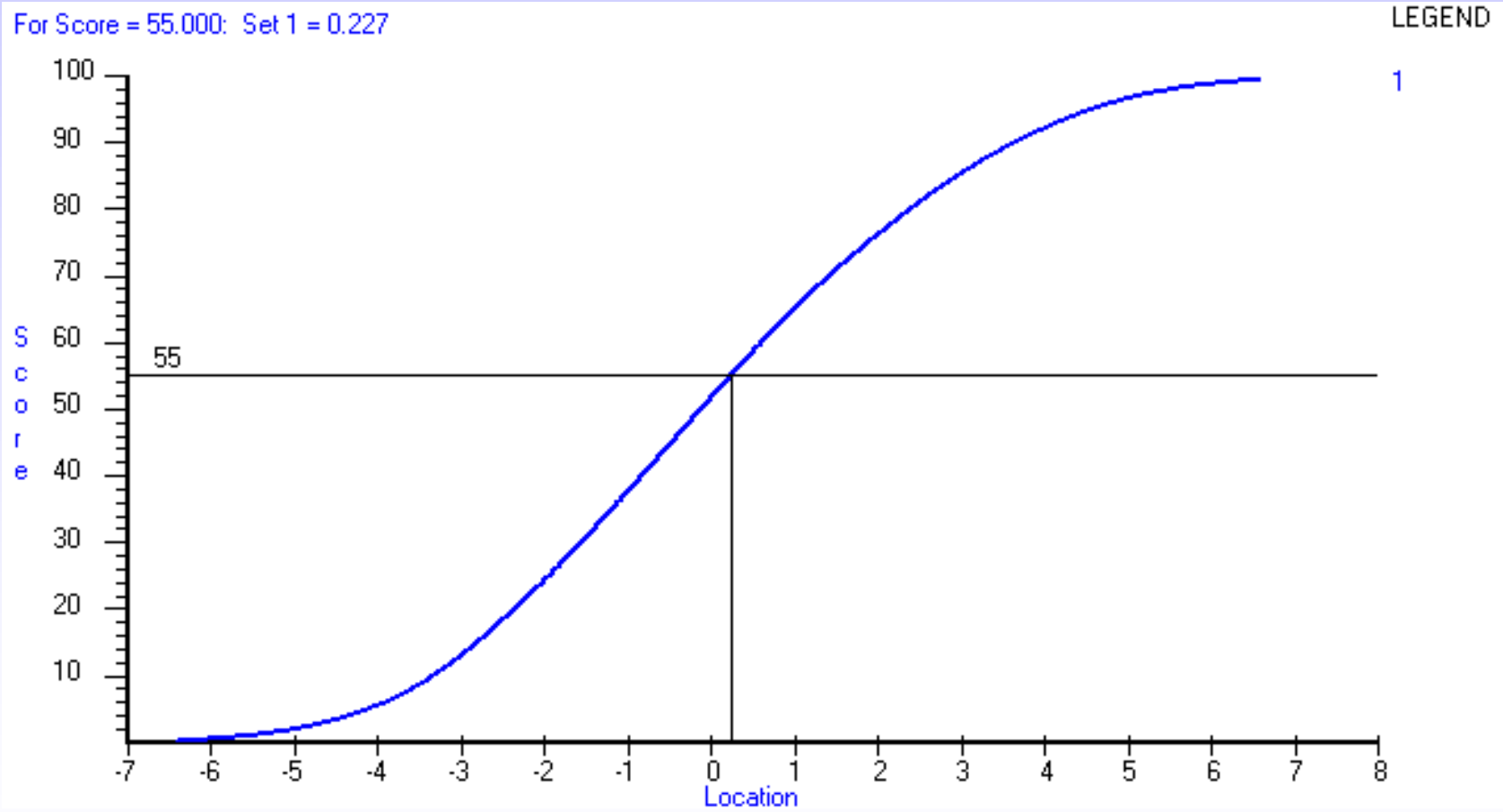
MID of - 5?


For Score = 93.000: Set 1 = 4.136
LEGEND
1
93
100
90
80
70
60
50
40
30
20
10
Score
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

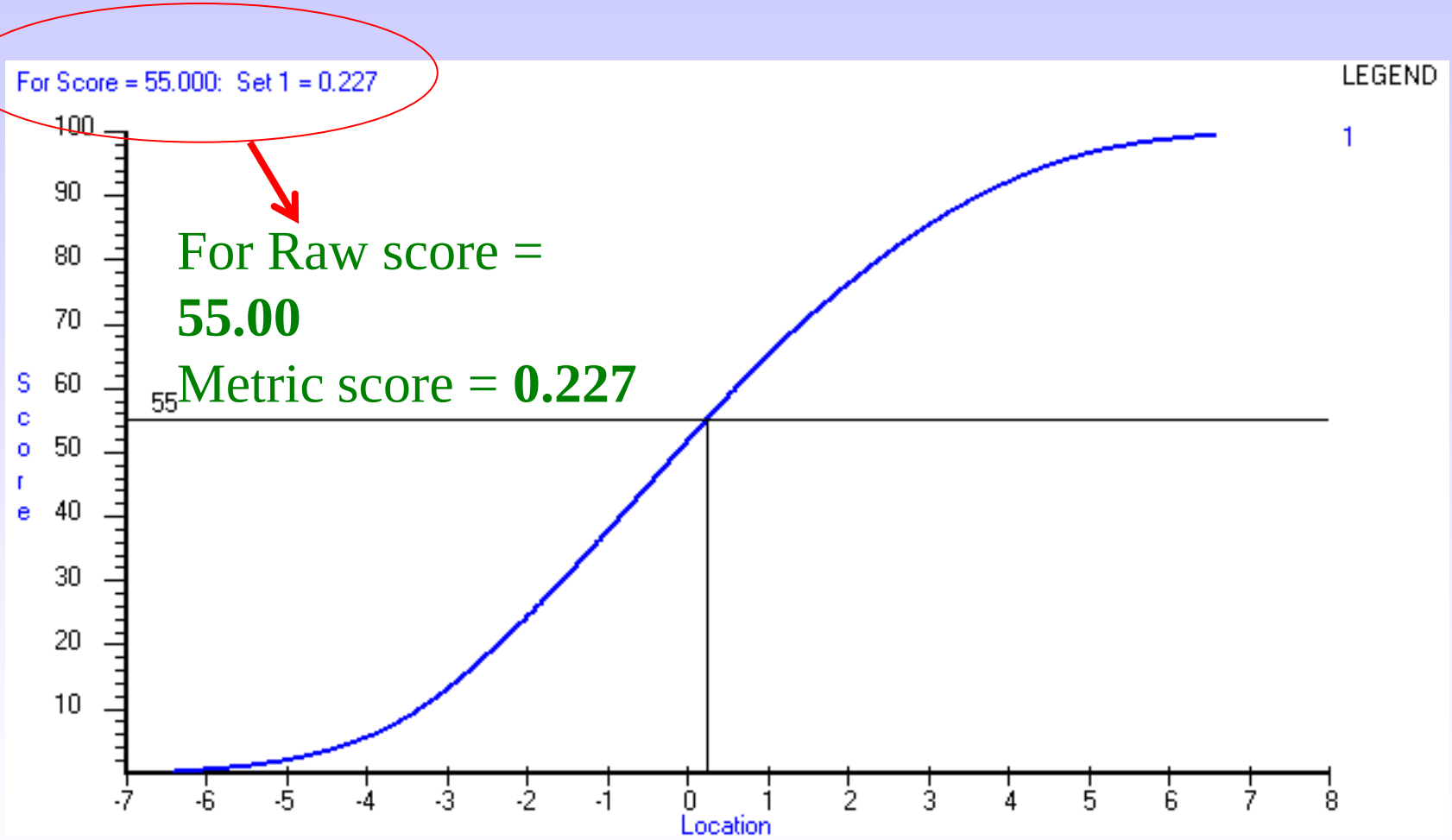
MID of - 5?


For Score = 93.000: Set 1 = 4.136
LEGEND
1
93
For Raw score =
93.00
Metric score = **4.136**
Raw Score difference of **-5** equates to a Metric Score difference of **1.38**
Score
Location

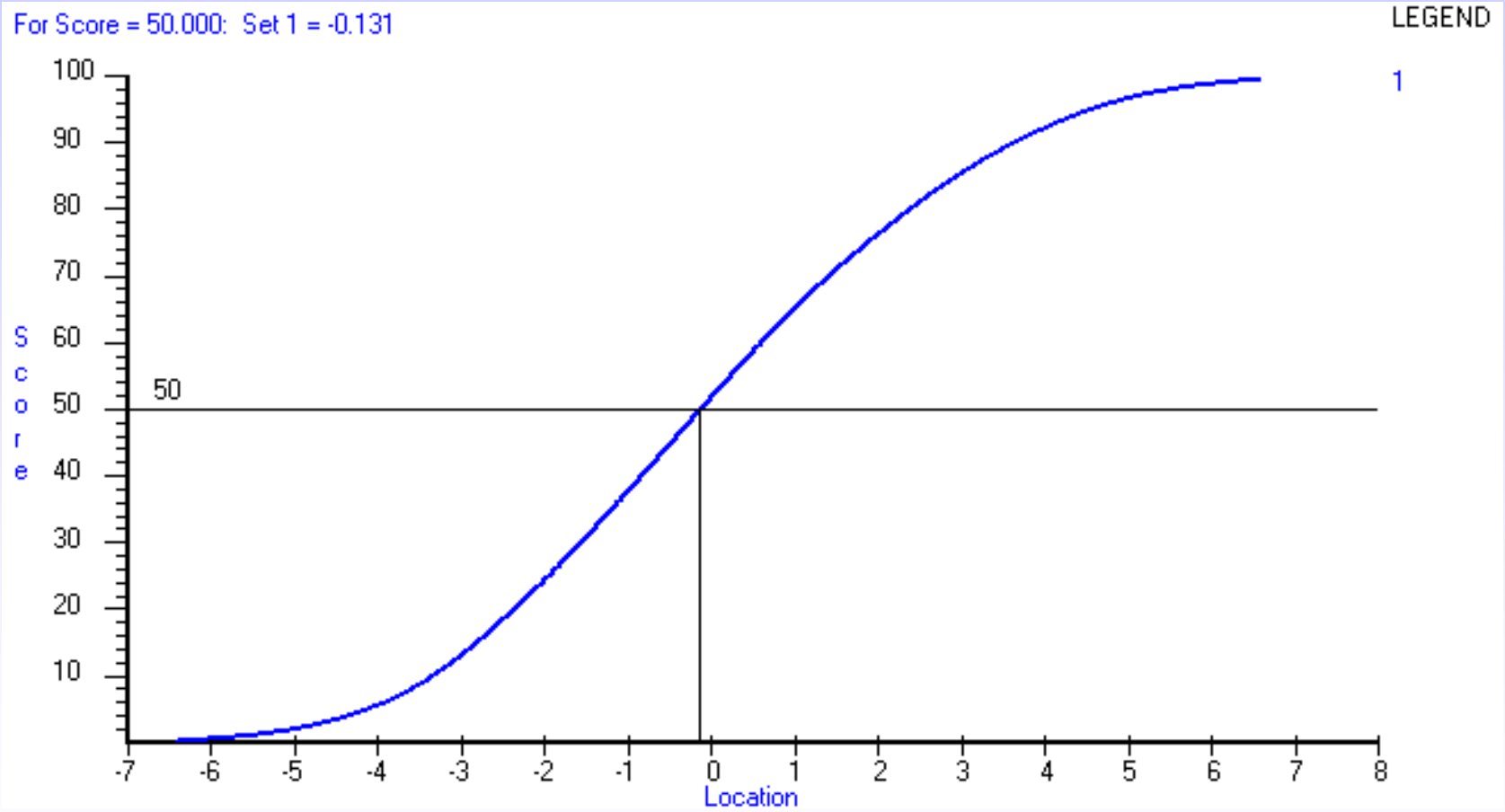
MID of - 5?


For Score = 55.000: Set 1 = 0.227
LEGEND
1
100
90
80
70
60
50
40
30
20
10
Score
55
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

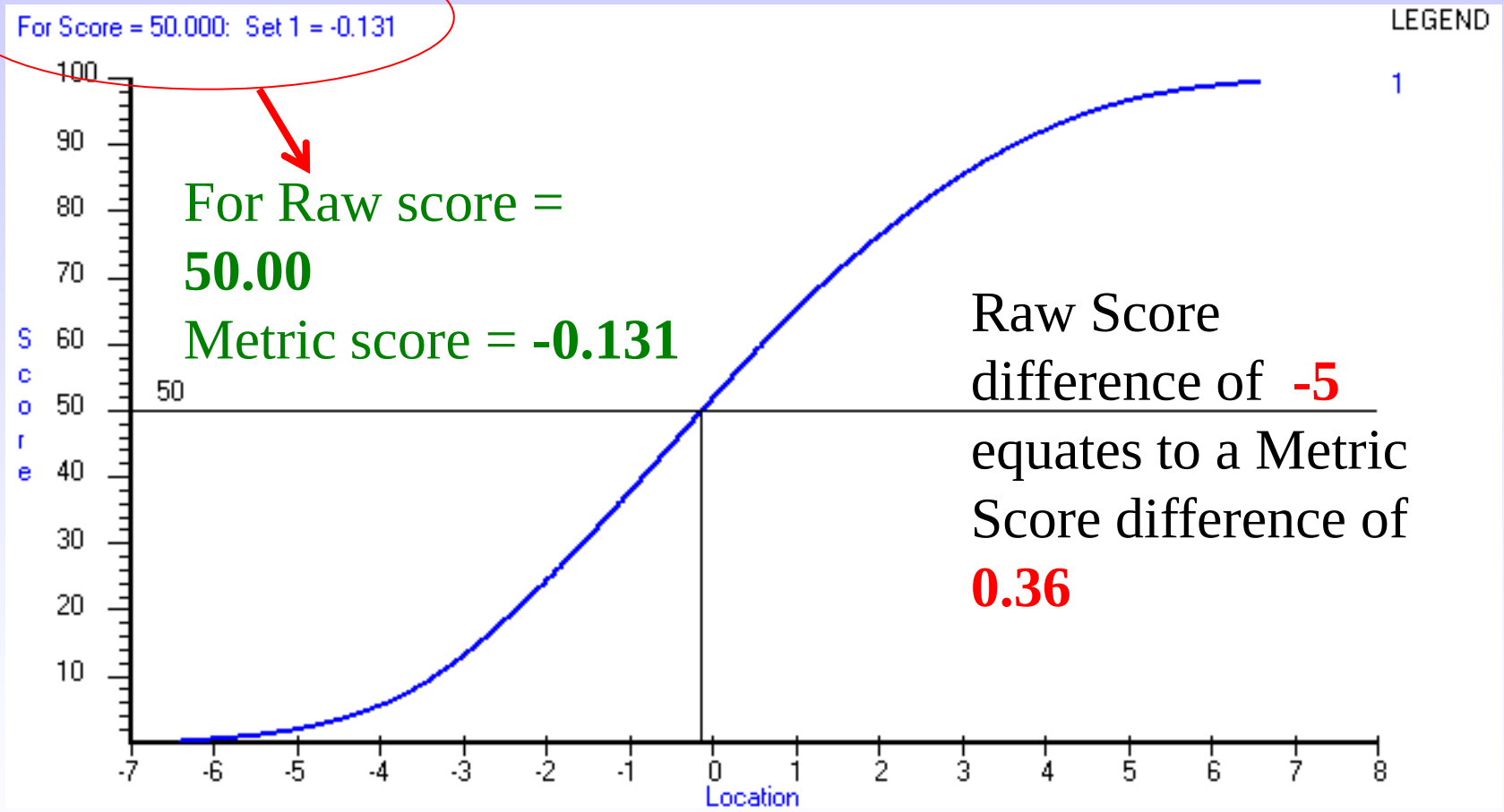
MID of - 5?


For Score = 55.000: Set 1 = 0.227
LEGEND
1
For Raw score =
55.00
Metric score = 0.227
55
Score
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

MID of - 5?

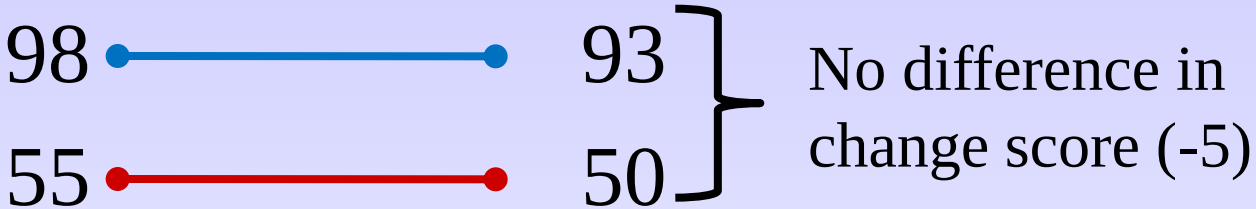

For Score = 50.000: Set 1 = -0.131
LEGEND
1
100
90
80
70
60
50
40
30
20
10
Score
50
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Location

MID of - 5?

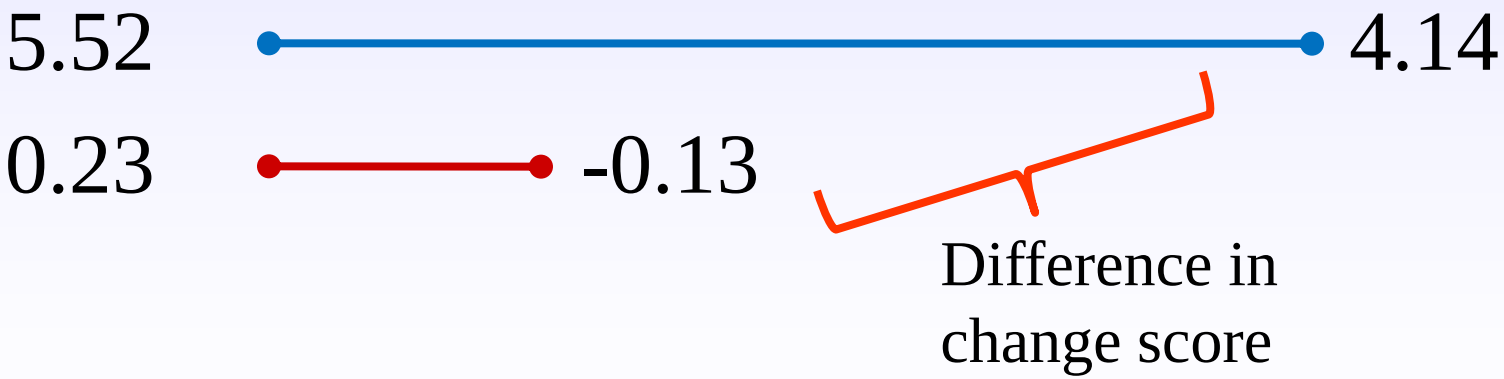

For Score = 50.000: Set 1 = -0.131
LEGEND
1
For Raw score =
50.00
Metric score = -0.131
Raw Score difference of -5 equates to a Metric Score difference of 0.36
Score
Location
100
90
80
70
60
50
40
30
20
10
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8

Minimally Important Difference?

With Raw Ordinal Scale Scoring:



With Metric Scale Scoring (drawn to scale)



MRC Clinical Trials Methodology Conference
October 4-5, 2011; Bristol, UK.

UNIVERSITY OF LEEDS

Can we use Ordinal Scales in this Way?

MRC Clinical Trials Methodology Conference
October 4-5, 2011; Bristol, UK.

UNIVERSITY OF LEEDS

Misinference from Ordinal Scales

- Using ordinal-level scores obtained with Likert-type scales to track change in clinical outcomes over time can result in erroneous conclusions. Comparisons of patient improvement by enrolment status are particularly vulnerable to the ambiguity inherent in employing change scores from an ordinal-level scale.

Using ordinal scales to track change over time: does the scale-level really matter?
Cook KF, Rabeneck L, Wray NP; Association for Health Services Research. Meeting. *Abstr Book Assoc Health Serv Res Meet.* 1998; 15: 264.

In Conclusion…

- There can be no such thing as a MCID, MID, effect size, or any other mathematical calculation on ordinal data
- Ordinal data are restricted to a change in magnitude – improved or deteriorated
- A considerable range of non-parametric statistical tests developed (many ranked based) to cope with this limitation
- Alternatively, where possible, data can be converted to interval level via correct application of the Rasch Model

Thank you for your Attention

MRC Clinical Trials Methodology Conference
October 4-5, 2011; Bristol, UK.

UNIVERSITY OF LEEDS